Training, Organizational Learning and Productivity: Three Essays on the Bangladeshi Garment Industry

by

Andreas Menzel

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Economics

September 2015

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

I am indebted to my thesis advisers Christopher Woodruff and Rocco Macchiavello for their guidance and support throughout my Ph.D., and for opening the possibility for me to do my own field work in Bangladesh. I would furthermore like to thank Eric Verhoogen, Emily Breza and Jonas Hjort for many useful comments while spending a term at Columbia University as visiting graduate student. I also want to extend my thanks to everyone who provided comments during and after the presentation of my work at seminars in Warwick, Columbia University, and various workshops that I could attend. Finally, I want to thank the University of Warwick Economics Department's community of Ph.D. students for making the time of my graduate studies highly rewarding and enjoyable, and my parents without whose support this thesis would not have been possible.

Declarations

All three chapters of this thesis contain original research based on novel data used for the first time in this research. Chapter three and four have been written solely by me, incorporating many comments from my thesis supervisors and others with whom I discussed my research. Chapter two is the result of a joint research project with my thesis supervisors and Atonu Rabbani from University of Dhaka, Bangladesh, in which we jointly planned and implemented the field work and analysed the data. All errors in the thesis are mine.

Abstract

This thesis consists of three main chapters, which address different but related research questions, using original data collected during extensive field work in the Bangladeshi garment industry.

After the introduction, Chapter 2 addresses possible reasons for the low share of women in supervisory positions in the Bangladeshi garment sector. Despite women making up 80% of the workers in the sector, they hold less than 10% of supervisory positions. Together with local partners, we designed a randomized intervention in which we trained equal numbers of male and female workers for supervisory positions, and placed them as supervisors on randomly selected sewing lines in their factories. Initially, lines with male trainees showed higher productivity, though this difference vanished after two months. Surveys of workers in the factories show that workers on all levels regard women as lacking the technical expertise to be good supervisors, while their leadership and other soft skills are regarded more favourably. However, extensive knowledge testing revealed that women have no less technical expertise, while management exercises and especially self rated ability revealed that women lack confidence and leadership skills compared to their male peers. This points to a mismatch between perceived and actual weaknesses of women as supervisors in that industry, which could prevent the management from taking effective measures to bring more women into supervisor roles.

Chapter 3 studies the effect of knowledge exchange among line supervisors in these factories on productivity. Specifically, it addresses the wide spread practice in economics to measure learning among co-workers through productivity increases, which, however, could also be caused by other peer effects, such as competition or imitation. I show that similar productivity increases as commonly used as evidence for learning are prevalent in situations in which learning is unlikely. However, a randomized communication intervention implemented by the respective factory management at three factories shows that knowledge exchange on production processes among workers indeed increases the efficiency of workers. There is furthermore some evidence that this effect was stronger between socially connected workers.

This effect of social ties in the communication intervention was based on social network data collected among supervisors in four garment factories. Chapter four discusses this network data in more detail, thereby contributing to several ongoing debates in network research.

Chapter 1

Introduction

There has been a recent surge in interest in development economics into studying firms in low income countries, mainly spurred by research showing that these firms exhibit widely differing productivity levels even within narrowly defined sectors, which could also explain the overall lower productivity in these economies (Hsieh and Klenow [2009]; Banerjee and Duflo [2005]). Furthermore, many large firms in developing countries do not implement simple management techniques which could greatly enhance their productivity (Bloom et al. [2013]). This could be due to low competitive pressure, or too burdensome regulatory environments and low levels of generalized trust which makes firms reluctant to involve managers from outside the founder family who could introduce better management techniques (Bloom et al. [2013]; Bandiera et al. [2015]). Unfavourable pay-schemes, which disincentivise workers to cooperate in the implementation of productivity increasing new technologies remain in place due to a lack of managerial expertise in many of these firms (Atkin et al. [2015]). Furthermore, ethnic and other kind of conflicts in many low income countries directly play out in these firms and affect their productivity (Hjort [2014]; Ksoll et al. [2010]; Kato and Shu [2011]).

This thesis contributes to the rapidly expanding body of research on large firms in developing countries in the context of the Bangladeshi garment industry. Bangladesh has over the last years emerged as the third largest garment exporter in the world,[1] and its vast garment sector, consisting of more than 5,000 large export oriented factories, provides a unique laboratory to study the opportunities and challenges that large firms in developing countries face (IGC [2014]). Furthermore, these factories, or at least the large share among them which is specialized in knit

[1] Source: WTO, International Trade Statistics 2014: www.wto.org/english/res_e/statis_e/its_e.htm

and woven garments, are remarkably homogeneous in their internal operations. All factories consist of at least three production departments: cutting, sewing, and finishing. 50-80% of the workforce is employed in the sewing departments, which are usually organized into parallel sewing lines, each designed so that the whole sewing process of a garment can be completed on one line. The sewing lines can therefore be considered as independent production units under the roof of one factory. Also, the sewing workforce is remarkably homogeneous across the knit and woven factories, being around 80% female and between 18 to 30 years old. Unless promoted to supervisory, mechanic, or quality control positions, which overwhelmingly happens only to males, workers typically drop out of the sector in their 30s (Chapter 2 of this thesis studies in detail the reasons for the extremely low promotion rates of female workers to higher positions in these factories). And even though factories vary widely with respect to their age, professionalism of their management, and adherence to regulations, this is not reflected in the pay of their ordinary workers, which generally follows government set levels for workers 'grades', such as 'helper', 'machine operator', or 'multi-task machine operator'.[2] Anecdotal evidence also shows that factories are not differentiated in which type of workers they hire, such as already experienced workers, or older or younger ones. Due to high worker turnover, factories constantly need to hire new workers, and usually train the new hires themselves for the required tasks. These characteristics lead to a setting in which a large number of factories with very similar organizational set-up and workforce, but considerable variation in management techniques, exist in a geographically small area. This presents a unique setting to study the interplay between management and productivity in a development context.

This thesis consists of three chapters on different but related research questions, using original data collected in extensive field work at several dozen garment factories in Bangladesh. The first chapter, written jointly with Christopher Woodruff, Rocco Macchiavello, and Atonu Rabbani, addresses the strong gender imbalance when it comes to supervisory positions in the garment factories. While 80% of the sewing workforce in the factories on average is female, women make up less than 10% on even the lowest supervisory positions. This question has potentially wide reaching implications, since if female supervisors were actually better at directing the overwhelmingly female workforce, the factories would forgo possible productivity increases by sticking with male supervisor, in a sector which is subject

[2]This relates primarily to the *nominal* pay of workers, according to their contract. Factories could still differ in the reliability with which they pay wages, or the arbitrariness with which they deduct sums from the wages for various reasons.

to strong price pressure and international competition. We gave to equal numbers of male and female workers, who were selected by factories as possible candidates for promotion to supervisory positions, a six week long intensive training program designed by the German bilateral development agency (GIZ), and subsequently placed them for a two months trial as assistant supervisors on randomly selected lines. During the trial phase, lines receiving a male trainee seemed to profit in terms of higher productivity, while those receiving female trainees did not, with the difference being statistically significant. However, this difference vanished in the subsequent months, with the lines which received female trainees catching up in terms of productivity. We accompanied the intervention with extensive surveys of workers on different levels in the factories, which revealed a number of additional insights. Workers in general rate men as more able supervisors. This gap was mainly driven by a perceived advantage of male supervisors in technical knowledge about machines and production processes, while women were rated little worse compared to men in leadership and communication skills. However, in tests on technical knowledge about garment production conducted at the beginning of the training session, no difference in technical expertise between female and male trainees could be found. On the other hand, men have much more confidence in their ability as supervisors compared to women, and women do somewhat worse at leadership exercises. However, these differences disappeared in later survey rounds after the training and trial as supervisor. We also found that the perceptions about female supervisors, especially among male workers, do improve if the worker has actually worked under a female supervisor. Taking these results together, we think they point towards a mismatch in perceived and actual weaknesses of women as supervisors in the local garment industry, shared by workers and managers. Women are perceived to lack technical knowledge for the supervisor job, while in fact the evidence points towards a lack in confidence and leadership skills, which however, seems to fade quickly with some training and experience. However, wrong perceptions about the mismatch could prevent the management to take effective action to increase the numbers of female supervisors.

The second chapter addresses organizational learning in these factories, or the question to which extend workers can profit from production knowledge gained by their co-workers in the factory. Organizational learning has long been assumed to be a key driver of productivity growth in firms (Arrow [1962]; Lucas [1993]). However, the study of the effect of knowledge exchange on worker productivity is difficult, as such knowledge exchange is usually difficult to observe. The literature often measures learning by increases in productivity if others have already worked on the same product. However, this risks confounding the effect of knowledge exchange with other peer effects, such as competition. I demonstrate this problem by showing that in Bangladeshi garment factories, sewing lines are more productive in the first days they produce a new garment, the more of the garment has already been produced on other lines before. However, similar effects on productivity can also be found among first lines that produce a garment which has not yet been produced on any other line before, if more than one line start producing the garment on the same day. As no other lines have produced this garment before in these situations, learning effects are unlikely to drive these increases. To clarify to what extent knowledge exchange between workers contributes to these productivity increases, a communication intervention was implemented at three Bangladeshi garment factories. On randomly selected sewing floors, whenever a sewing line started to produce a new garment which had already been produced on another line before, the line chief from the line that already produced it was sent by the production manager to brief the line chief who now also started producing the garment. I show that these briefings increased productivity of the sewing lines on the first two days they produced the new garment, before the line would reach its long run productivity levels again. This provides novel experimental evidence that knowledge exchange indeed drives productivity increase in firms. There is furthermore some evidence that this effect was stronger when the line chief reported social ties to the line chief who provided the briefing.

The result on the interaction of the randomized communication intervention with social ties is based on survey data on social connections among line chiefs, which I collected in four Bangladeshi garment factories. While Chapter 3 studies one of the possible effects of these social ties, increasing knowledge exchange between socially connected worker, Chapter 4 discusses the characteristics of these social networks in more detail and tests to what extent several network formation models can explain these characteristics. The line chief networks exhibit low levels of density and high levels of clustering, a common characteristics of empirical social networks, which many network formation models struggle to replicate. I show that in the context of my data, a simple block random graph model, with blocks defined on the level of the sewing floors in the factories, does capture both density and clustering levels surprisingly well. The high levels of clustering in the network data could be driven by line chiefs forming ties within small spatial units. This suggests that the ubiquitous high levels of clustering observed in empirical social networks could also more generally be caused by social ties being formed in very localized interactions. In a further contribution, the chapter shows that lagged in-degree of line chiefs predicts to what extend newly arriving line chiefs form social ties with existing line chiefs. While this 'rich-get-richer' phenomenon is interesting in its own right, this results also confirms a central underlying assumption of the popular preferential attachment random graph model (Barabasi and Albert [1999]; Jackson [2008]).

To conclude, this thesis uses the unique setting of the Bangladeshi garment sector, which consists of a large number of fairly homogeneous large firms in close spatial proximity, to contribute to several research questions. It provides suggestive evidence that wrong perceptions about the strengths and weaknesses of female supervisors hold back women from advancing to roles with more responsibilities and pay. It furthermore shows that communication of production knowledge among workers on the same level in the factory hierarchy spurs productivity, and contributes to the understanding of social network formation in the workplace. The following three chapters will discuss this work in more detail.

Chapter 2

Challenges of Change: An Experiment Training Women to Manage in the Bangladeshi Garment Sector

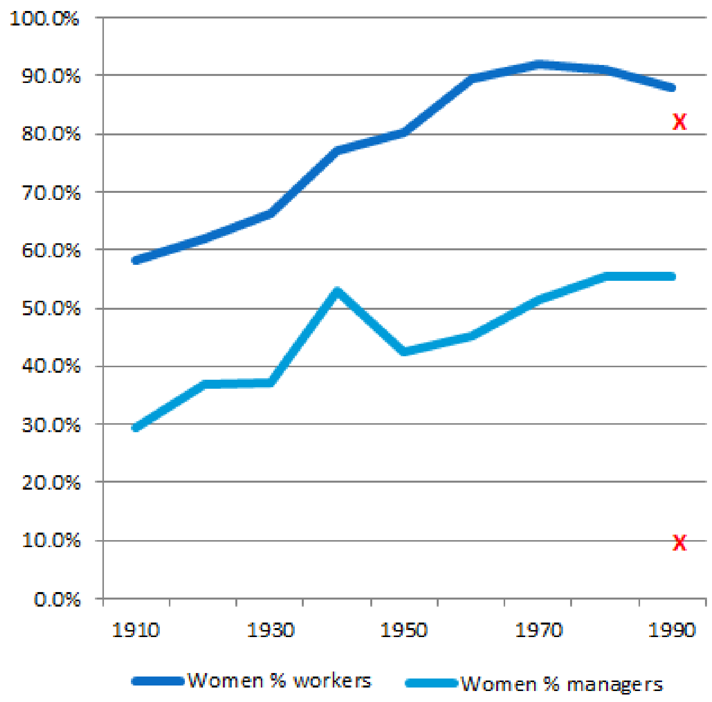
Joint with Rocco Macchiavello[1], Atonu Rabbani[2], and Christopher Woodruff[3]

2.1 Introduction

Management of large firms in low-income countries is highly variable and poor on average (Bloom et al. [2012]). While the recent literature has focused on the broad set of management practices pioneered by Bloom and Reenen [2007], effective management – including the adoption of such practices – rests on successfully managing relationships and perceptions in the workplace (Gibbons and Henderson [2012]). This observation shifts our attention from practices to managers. With shortages of qualified managers perceived to be an important barrier to better management in developing countries (McKinsey [2011]), we still know little about how companies in low-income countries develop and select managerial talent.

[1]University of Warwick
[2]University of Dhaka
[3]University of Warwick

Figure 2.1: Female Worker and Manager in the Garment Industry, U.S. vs. Bangladesh



Note: Figure shows the historical evolution of the share of female workers and managers in the US garment industry, and compares it against the current shares in the Bangladeshi industry. US Data from Ruggles et al. [2010], Bangladeshi data own calculations.

We study mid-level management in the ready-made garment industry in Bangladesh, a sector with more than 4,000 factories, employing around 4 million workers and accounting for an estimated 12 percent of Bangladesh's GDP. Besides its intrinsic relevance, the sector provides an ideal context to study mid-level managers. The sewing section in a typical factory is organized along several production lines employing between 20 and 80 workers (operators) managed by mid-level managers (line supervisors). We focus on one distinctive feature of the industry: while women account for about 75 to 80 percent of workers in the sewing operations, men account for around 95 percent of supervisors and higher-level managers. The situation is stark: Figure 2.1 contrasts employment patterns in Bangladesh with the historical evolution in the United States and shows just how strong the gender imbalance is in Bangladesh.

Why are there so few female supervisors? Does this gender imbalance result in a large misallocation of managerial talent in the sector? To address these questions, we start from a simple observation: in a static sense, managerial capital is misallocated if the marginal female supervisor is more effective than the marginal male supervisor.[4] If this was the case, factories could improve efficiency by promoting additional women and fewer men.[5]

Empirically, we face several challenges. First, given there are so few female supervisors to begin with, it is difficult to identify the marginal female supervisor. To overcome this problem, we implement a six-week operator-to-supervisor training program in 24 factories.[6] The program induces factories to try out (and possibly promote) more female supervisors than they otherwise would. Second, we need to observe the performance of both male and female line supervisors. We implement an experimental design in which upon returning from training, trainees are tried as assistant supervisors on randomly assigned production lines. This allows us to identify the causal impact of having a female supervisor on performance. We then compare the performance of females and males trained in the program, and the response of operators working for them, using both very detailed production data and in-factory surveys. Finally, we recognize that the initial lack of female supervisors poses additional interpretative challenges to the test outlined above. We implement uniquely detailed baseline surveys and diagnostics tools with workers and managers at all levels in the factories to understand what supervisors are expected to do and compare perceptions and reality of women's relative abilities within the relevant pool.

We show four sets of results. First, we ask what supervisors (are supposed to) do, and what are the perceived weaknesses of females as supervisors. Across

[4] Note that the observation is correct for any distribution of potential supervisor's effectiveness across genders. In particular, it is possible that in the current industry equilibrium men self-select and/or invest in additional skills with the expectation of becoming supervisors. This could result in the pool of men available for promotion being on average better than the pool of available women for promotion.

[5] Large inefficiencies would be at odds with the fact that all factories in our sample are large exporters operating in highly competitive product markets. A large literature shows that competition increases efficiency (Syverson; Foster et al. [2008]; Backus [2014]), improves management practices (Bloom and Reenen [2007]; Bloom et al. [2012]) and that export status is associated with higher productivity (Bernard et al. [2007]) and better management (Bloom et al. [2015]). On the other hand the factories in our setting are typically owned by a small group of investors and might face lower pressure on the financial market side.

[6] The training program was designed by the German bilateral aid agency, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), together with local training companies.

eight broadly defined sets of tasks, we find remarkable agreement across hierarchical layers in the factories about what supervisors are supposed to be doing. There is also remarkable agreement in the factory that women are weaker than men in essentially all eight dimensions. In particular, women are perceived to be less competent than man in understanding machines and operations – crucially, the most important task for a supervisor from the point of view of operators. These negative perceptions are less strong, but nevertheless present, among female operators and among those operators with experience working under a female supervisor.

Second, we compare these perceptions to reality. Before the training began, we conducted an extensive skills assessment with the trainees. Three results emerge. First, there is no difference in technical knowledge of machines and operations between male and female trainees – despite the widely held opinion to the contrary. Second, in simple leadership exercises women are less likely to be selected by their team for a leadership position and women perform slightly worse in an exercise in which they instruct other team members to perform a simple task. Third, in essentially all eight broad tasks female rate themselves as being less good than existing supervisors while male trainees do not.[7]

Third, we examine the performance of male and female trainees once they return from the training. Two sets of results emerge. First, immediately upon returning from training female trainees underperform relative to male trainees. This initial gap in performance is measured both using surveys of operators supervised by the trainees as well as detailed daily line-level production data. The gap in performance, however, completely closes after few months working on the line as supervisors. In simulated management exercises, female trainees outperform male trainees on average but not when managing small teams that include a male operator.

Finally, we explore attitudes of male operators exposed to the program. These are of particular importance given that the bulk of future line supervisors is currently recruited from this pool. Two results stand out: first, male operators exposed to female trainees improve their view of female as supervisors (but less than female operators do). Second, male operators exposed to female trainees are more pessimistic about their prospects of being later promoted to supervisor roles and

[7]Relative to male trainees, female trainees also have lower education levels, numeracy skills and factory experience.

expect to work for a shorter period of time in the factory. In short, the promotion of female supervisor appears to demotivate male workers.

Taken all together, these results portray a nuanced but comprehensive picture of the causes and consequences of gender imbalance in the sector. The evidence is consistent with an industry equilibrium in which factories haven't experimented with female supervisors due to misperceptions about their relative effectiveness. The fact that misperceptions are widespread across the organization – including among workers and potential female supervisors themselves – supports this equilibrium by requiring simultaneous changes in beliefs. In a static sense, even a profit maximizing manager with correct beliefs might not promote women if - in our case - *he* believes other co-workers won't respond adequately due to their beliefs. Dynamically, such a manager might believe workers' perceptions can be aligned to reality, but at the cost of alienating and demotivating male operators – from which the bulk of managerial talent is still likely to be supplied to the factory in the short-run. In the conclusions, we distil some implications of this interpretation for our understanding of organization's failure to adopt adequate management practices, the sources and consequences of gender imbalances in general, and the design of policies that could ameliorate those.

This paper contributes to different strands of literature. It complements a literature examining the causes and consequences of the (lack of) female leadership. Although there are numerous contributions studying the gender gap in labour markets and in the private sector (see, e.g., Bertrand et al. [2014]; Matsa and Miller [2013]; Bertrand and Hallock [2000]; Dezsö and Ross [2012]; Glover et al. [2015]), our work is conceptually closer to studies of female politicians in India by Chattopadhyay and Duflo [2004] and Beaman et al. [2009].[8]

As Chattopadhyay and Duflo [2004] we focus on establishing the causal impact of female leaderships on outcomes. As Beaman et al. [2009] we emphasize the importance and evolution of perceptions of female leaderships. Our analysis, however, needs to be adapted to reflect the operations and incentives of large firms operating in a competitive export sector. First, the performance - not just the appointment - of female leaders is affected by beliefs and perceptions of co-workers. Second, we investigate the costs associated with appointing female leaders.

[8]Some of our results are also related to a large experimental literature documenting gender differences in attitudes and preferences, see, e.g., Gneezy and Rustichini [2004]; Niederle and Vesterlund [2007]; Niederle et al. [2013].

In so doing, the paper also contributes to the literature on management and productivity (see, e.g., Hsieh and Klenow [2009]; Bloom and Reenen [2007]; Bloom et al. [2012, 2013]; Bruhn et al. [2012]; McKenzie and Woodruff [2015]).[9] The work by Bloom and various co-authors raises a puzzle: the management practices they study are well-known and seemingly simple to implement. Why do firms fail to implement them? Gibbons and Henderson [2012] argue that changing practices is actually quite complex, both because individual practices are complementary to one another (see also Ichniowski et al. [1997]) and because management involves both formal rules and informal norms. Managers may know what is wrong, know how to fix what is wrong, but yet be unable to implement the required changes because they are unable to shift the equilibrium of the game between managers and workers (or between managers at different levels of the hierarchy). Our research design and emphasis on understanding misalignment of perceptions within the firm borrows from this perspective. The difficulties of implementing change echo recent work by Atkin et al. [2015] in the soccer ball industry in Pakistan. They document workers resisting to a new technology. We highlight how resistance to change is embedded into a set of norms and perceptions we set out to measure.

Finally, the paper contributes to our understanding of the garment sector in Bangladesh and elsewhere. Historically, the sector has represented one of the first opportunities for women to enter the formal labor force. Heath and Mobarak [2015] and Atkin [2009] study the relationship between garments, female labour force participation and schooling in Bangladesh and Mexico respectively. Mid-level managers in the industry are also studied by Schoar [2011] and Achyuta et al. [2014] with different focus and research design.

[9] There are two additional methodological contributions of the paper. With respect to the productivity literature, the paper uses a physical measure of productivity in a multi-product industry with product differentiation. Line-level productivity is measured taking advantage of "standard minute values" which allow to convert units of differentiated garment pieces into standardized measures of time value of output. With respect to the literature on the evaluation of training program, we directly investigate the impact of the training on productivity, not just on the wages paid to trainees. This is important as for a variety of reasons wages might fail to reflect the marginal value of labour.

2.2 Design and Data

The training program we implement was designed by GIZ, the German bilateral aid agency, in conjunction with local training companies. GIZ's goal in developing the program was to increase the number of women working as supervisors in the sector. The training was viewed as important to build skills of female operators, and to convince factories that women were equipped to be supervisors. The training lasts six weeks, with eight-hour sessions held at the classrooms at the training providers offices on six days per week. The curriculum was divided more or less equally into modules on production planning and technical knowledge, quality control, and leadership and social compliance. We initially contracted with three training providers and then later selected one of them with the capacity to conduct all of the sessions.

The project was carried out in two phases. Phase 1 began in November 2011 and continued through February 2013, with 56 factories sending five participants each to training. After analysing the data from the first phase, we made several changes to the project design and launched the second phase in February 2014. Lessons from the first phase were incorporated into the design of the second phase. As a result of incorporating the initial lessons, the quality of the data are generally higher in the second phase. Aside from a management simulation exercise that we conducted only in Phase 1, we rely on the data from the second phase in the analysis. We describe the design for the second phase here, and refer the reader to Appendix A for a description of the design of the first phase.

In the second phase of the project we worked with direct and indirect suppliers of a large UK-based buyer. We started with a pool of 26 suppliers of woven and light-knit products located in the Dhaka area.[10] The buyer invited these suppliers to an information session in February 2014. At the end of the information session, 24 factories expressed interest in the project, all of whom ultimately participated.[11]

We asked each factory to consider the expected demand for new supervisors in the factory in the months following training, and to select a number of trainees matching that demand. Because the size of the factories varied and because, for

[10]We limited the sample to the Dhaka area for logistical reasons and to woven and light-knit because production in these products is organized by sewing lines in Bangladeshi factories. Direct suppliers are managed by employees working directly for the buyer; indirect suppliers are managed on behalf of the buyer by intermediaries.

[11]Five of the factories sent operators to the first training session, but dropped out in the second half of the program.

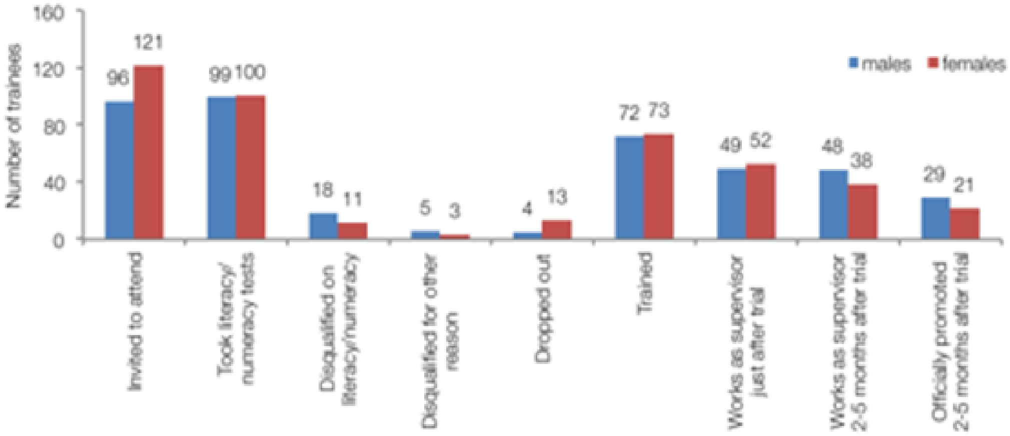
example, some factories were planning to open new production lines, the number of trainees varied from as few as four to as many as 24. Where an even number of trainees was provided, we asked factories to select an equal number of male and female trainees. Where an odd number of trainees was selected, we asked them to select one more female than male. We informed the factories that much of the training material was written, and therefore the trainees needed to have at least basic literacy skills. We gave them no other criteria, but did encourage them to involve managers down to at least the line chief level in the decisions. The factories sent 99 males and 100 female trainees to the training centre for the initial diagnostic. Note that this represents a significant push toward female supervisors, as in the typical factory only around 4 percent of supervisors were female at baseline.

We scheduled four training sessions, the first beginning March 9^{th}, 2014 and the last beginning June 1^{st}, 2014. Each factory was randomly allocated to training rounds 1 and 3 or to training rounds 2 and 4, and the trainees from the factory were randomly assigned to receive early or late training. Randomization at the trainee level was stratified on gender so that a nearly equal number of female and male trainees were trained in each session.

Training sessions met six days per week for roughly eight hours per day, and the training lasted six weeks. Factories agreed to give each trainee a six- to eight-week trial as an assistant line supervisor immediately after the end of the training program. We asked factories to identify the lines which were suitable for the trainee trials and to identify an experienced supervisor working on each of those lines who could act as a mentor for the trainee. On the penultimate day of training, we invited the mentor supervisors to the training centre and matched them randomly with one of the trainees from their factory – thus assigning the trainee randomly to a production line for the trial period. On the day the mentors attended the training centre, we conducted a series of team building exercises between trainees and mentors. After the six- to eight-week trial, factories were free to return the trainee to a position as operator, leave them as an assistant supervisor, or promote them to supervisor.

There was dropout of trainees at various points, detailed in Figure 1. The factories initially selected 121 females and 96 males for training. All were invited to the training centre for the initial assessment. On the allocated day, 100 females and 99 males actually showed up. Twenty-one females declined to come to the training

Figure 2.2: Selection, Training, Trial and Promotion of Trainees



Note: Figure shows the number of female (red bars) and male (blue bars) trainees which participated or dropped out in different stages of the project, and which got promoted to supervisor levels in their factories.

centre, either because they decided they did not want to be supervisors or because of resistance from their families. Meanwhile, three additional males came as some factories replaced the females who declined to attend. Admission to the full training program depended on passing the literacy and numeracy test administered at the training centre. The literacy exam was developed in conjunction with researchers at BRAC University.[12] Nominees were disqualified if they scored zero on either the literacy or numeracy exam, or if they scored below 25 percent on both parts of the exam. Eleven females and 18 males did not pass the literacy / numeracy threshold. An additional three females and five males were disqualified for other reasons, mainly either because the factory sent a male rather than a female.[13] Finally, after the assessment day, 13 females and four males decided they did not want to complete training and dropped out of the program. The remaining sample, all of whom completed the training course, was 73 females and 72 males. Figure 1 also shows the number of trainees working as a supervisor at various points after training, which we discuss in more detail below.

[12]The literacy/numeracy test was developed by Sameeo Sheesh and Badrul Alam of BRAC University's Institute of Education Development (IED). The content is based on the skills required to benefit from the Operator to Supervisor Training material, and content taught in grades 5 through 8.

[13]In a couple of cases, the literacy exam was mismarked so that a failing score was given when the exam was a marginal pass.

2.2.1 Data

We conducted surveys on six separate occasions. Prior to the start of training, but after factories nominated the trainees, we conducted a baseline survey at the factory and a combined survey and skills assessment for the trainees at the training centre, which took a full day. In addition to gathering basic information on demographics, work history and attitudes, we assessed knowledge of skills and production processes, conducted communication, teaching and leadership exercises, and tested numeracy, literacy and non-verbal reasoning skills. The assessment is described in more detail below.

Near the end of the six-week training program, we asked factories to nominate a number of production lines equal to the number of trainees where the trainees would work as assistant supervisors for a period of at least six to eight weeks. We also asked the factory to nominate one existing supervisor from each of these lines who would serve as a mentor for the trainee. We then randomly assigned the trainees to one of the nominated lines / mentors. With the list of lines and mentors in hand, we conducted a baseline survey in the factory just prior to the start of the trail. For the factory survey, we surveyed line operators, line supervisors, line chiefs, assistant production managers, production managers and HR managers. Three operators and all of the supervisors and line chiefs were surveyed at the lines where trainees would have their trial. Line chiefs from the lines where trainees were working at the start of the training were also surveyed. The three operators were randomly selected from the line in a way which ensured that at least two of these operators work directly under the mentor supervisor, and that we select both male and female operators wherever possible.

On the last day of the training, the mentors were invited to the training centre and paired with their matched trainee. We conducted team building exercises and also conducted a follow-up survey and skills assessment. The purpose of this survey and assessment was to capture any effects of the training on trainees, and to measure the skills of experienced mentor supervisors for comparative purposes. At the end of the six- to eight-week trail period, we again invited the trainees back to the training centre for refresher sessions and group discussions of their experience during the trial. We also conducted a final skills assessment for trainees to measure the effect of the factory trial.

The fifth survey was conducted in the factory after the trial period ended. We again surveyed three randomly selected operators, the supervisors and line chiefs of the lines that were nominated for the trial, and the assistant production managers, production managers, and HR managers. In addition, where there was either non-compliance with the assignment of trainees to lines, or where trainees had moved from the assigned line to another line after the trail began, we surveyed lines which were not nominated for the trial, but where trainees were actually working as assistant supervisors.

Finally, we conducted a second follow-up survey in the factory either two- or five months after the trial ended. The survey sample was selected using the same criteria as the previous factory survey. Operators and supervisors were surveyed from lines where a trainee was working as either an assistant line supervisor or a line supervisor. In addition, all of the trainees were surveyed in-person if they were still working at the same factory, and over the phone, if they had left.

In addition to the face-to-face surveys, we conducted telephone follow-up surveys with trainees at regular intervals. During the six- to eight-week trial, we contacted the trainees every week to track the line they were working on, and the level of responsibility given to them. We also asked the trainees to keep a daily diary of their experience working as an assistant supervisor or supervisor. After the trial ended, we contacted the trainees every month until March 2015 (four to nine months after the trial) to track where they were working, and their designation.

In addition to the survey data, we also collected daily line-level production data from each factory. We describe these data in more detail in Appendix B and in Section 2.6 below.

2.2.2 Characteristics of Trainees

Table 2.1 shows basic demographic and skills data for the pool of trainees, compared to existing supervisors and random operators where the comparison data are available. Compared with a sample of random operators, the trainees have two additional years of schooling and just more than half a year more tenure in the factory. Age, marital status and experience in the garment sector are similar to other operators. We split the supervisor sample into mentors and non-mentors for the purposes of comparing trainees with existing supervisors. We see that, while the trainees have much more schooling than typical operators, they have almost a year less schooling than typical supervisors. They are also 4.7 years younger with 2.3 years less experience in the sector. However, the age of the trainees is statistically identical to the age of the random supervisors at the time of their promotion to supervisor. With regard to the relative skills of female- and male trainees (not shown on table), we find that females are just over a year younger (p=0.05), but there are no differences in schooling or experience.

Whether the trainees have less schooling than existing supervisors because factories face a shortage of workers with higher schooling levels, or whether the factories have not selected the very best supervisory talent for the training program is not clear. But while 62 percent of existing supervisors have at least a lower secondary certificate (that is, they have passed O-level exams), only 14 of 430 random operators (3 percent) have achieved this level of education. This suggests that factories do face a very limited pool of workers with education levels comparable to the pool of existing supervisors. This, combined with the age and experience profiles of the trainees suggests that the factories selected trainees in a manner similar to those selected in the usual promotion routine.

We can also compare the skills of trainees and the mentor supervisor using tests administered at the training centre during the skills assessment, though we lack similar data for other operators and supervisors. The bottom half of Table 2.1 shows that the literacy and numeracy scores of the trainees are significantly below those of the mentor supervisors. These data provide further evidence that the skills of the trainees are below those of the mentor supervisors.

2.3 Perceptions of female supervisors

We asked employees at all levels of the factories to tell us which tasks are the most important for line supervisors. We first constructed a list of eight main tasks from an initial set of open-ended conversations with managers. We then gave each respondent 10 tokens and asked him or her to place the 10 tokens on the list of the eight tasks to indicate the relative importance of each. Respondents were told they could place all 10 tokens on a single task if they thought that it was the only one that is important, or spread the tokens across the tasks as they wished. Surveys were conducted with HR Managers, Production Managers, Assistant Production Managers, Line Chiefs, Line Supervisors and Operators.

Table 2.1: Demographic Characteristics

Panel A: Trainees vs. Operators and Supervisors

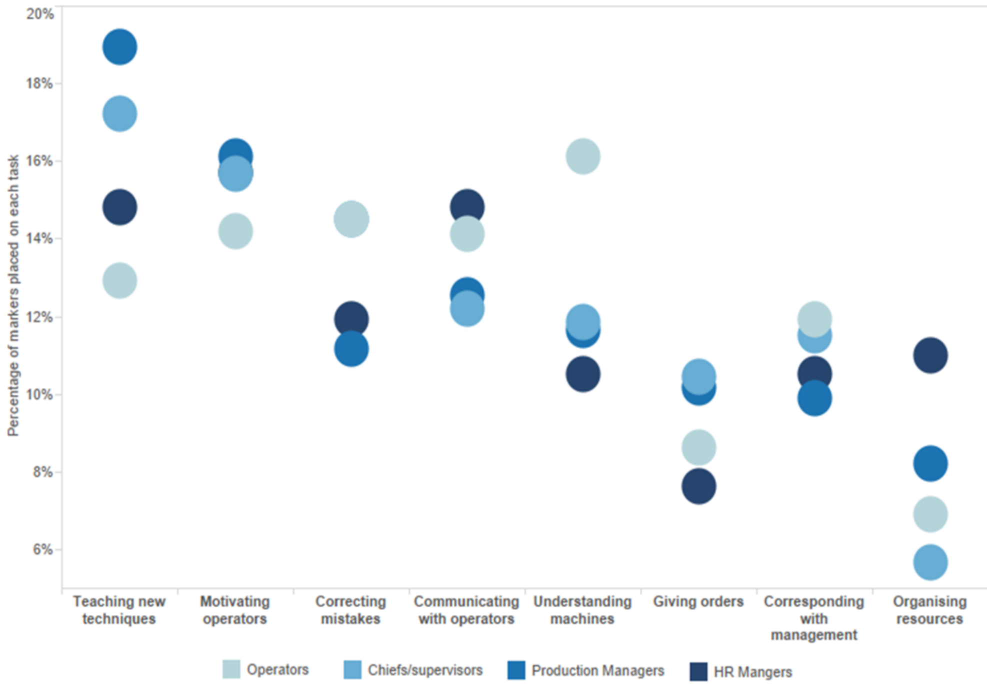
	Mean				Comparisons	
	Trainee Pool	Operators	Random SVs	Mentor SVs	Trainees vs. Operators	Trainees vs. Random SVs
	N = 199	N = 430	N = 92	N=142		
Gender (female =1)	0.50	0.73	0.04	0.04	-0.23***	0.46***
Age (current)	24.4	24.1	29.1	29.3	0.30	-4.66***
Age at promotion to SV	24.4	NA	25.3	24.3	NA	0.84
Married	0.71	0.77	0.85	0.89	0.06	-0.14**
Education (yrs in school)	7.83	5.80	8.77	9.56	2.03***	-0.95***
Working in Garments (yrs)	6.53	6.10	8.83	9.21	0.43	-2.30***
Tenure in Factory (yrs)	3.41	2.78	3.30	3.80	0.63***	0.11***

Panel B: Trainees vs. Mentor Supervisors

	Mean		
	Trainee Pool	Mentor SVs	Trainees vs. Mentor SVs
	N = 197	N=113	
Literacy	7.14	9.54	-2.39***
Numeracy	3.4	5.2	-1.72***
Non-verbal reasoning	2.75	2.78	-0.03

Notes: Table shows mean characteristics of trainees, random sewing operator from lines at which trainees worked before training, random line supervisors, and line supervisors selected by factories as mentor for trainees, on whose lines trainees will be trialed as supervisors after training. Statistical significance of differences in comparisons: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure 2.3: Tasks of Supervisors: Attached Importance



Notes: Workers on various levels in 26 factories were asked to place 10 tokens on a list of eight general tasks of line supervisors (generated after open ended conversations with several factory managers), according to the relative importance they attach to the task.

Figure 2.3 shows the percentage of tokens placed on each of the eight tasks by respondents holding different positions at the factory. The characteristics given the highest weights are shown to the left of the graph. One pattern that emerges is that there is very close agreement about which characteristics are important among all levels of managers. Teaching and motivating operators are given the highest weights by all managers. Operators, on the other hand, appear to prefer problem solvers, giving slightly higher weights to understanding machines and correcting mistakes. There is agreement across the hierarchy that organizing resources, corresponding with management, and giving order are less important tasks of supervisors.

We then asked the same set of respondents whether, based on their own experience, they thought females or males were better at each of the eight tasks of being a supervisor. The allowed responses included the option of "both are equal". We code these data in a way that indicates the perceived deficit that females face in each of the tasks. A response “males are better” is coded as -1, “females are better” is coded as +1 and “both are equal” is coded as 0. The scores are shown in Figure 2.4, again by type of respondent.[14] The first takeaway from the table is that males are overwhelmingly seen as having an advantage in every supervisory task. Line operators and line supervisors rate males better in all eight tasks, line chiefs and production managers rate males better in seven of the eight tasks, and HR managers see males as being better in five of them.

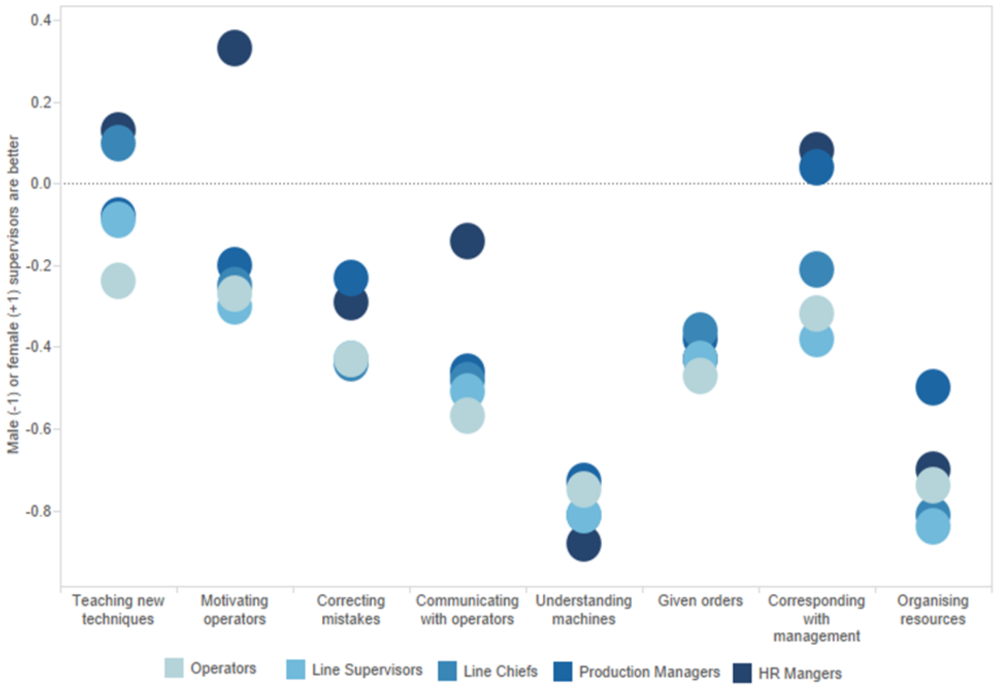
We also find a very high level of agreement about the specific tasks where females are most lacking. According to every category of respondent, females have the largest deficits in understanding machines and organizing resources. All respondents also agree that the three areas where females are closest to males are teaching new techniques, motivating operators, and corresponding with management, though there is some disagreement about the ranking of these three. Notice that the two tasks rated as most important by managers are two of those where the gap between females and males is perceived to be the smallest. On the other hand, machine knowledge, rated highest by operators, is the area where females are perceived to be the weakest.

The sample of operators is the largest and most diverse, so in Figure 2.5, we show the same comparisons for different subgroups of operators. First we split the randomly selected operators by gender. The relative rankings are very similar for female and male operators – the correlation is 0.87 – though female operators uniformly describe a smaller gap. Next we split the operators into those who have and those who have not worked for a female supervisor at some point in their career. Past experience working for a female supervisor has no significant effect on the perceived gap in female skills. Finally, when we asked the trainees the same comparisons between generic male and female supervisors, the responses are very close to those of other operators. As Figure 2.5 shows, female trainees do rate women somewhat higher than do other operators.

We also asked trainees about their own ability relative to typical supervisors in their factory. We first asked the trainees to rate the typical supervisor on a scale of 1-10 with regard to each of the eight supervisory roles, and then asked the trainee to rate her/him self on the same scale. Female trainees rate themselves as

[14]We did not ask the Assistant Production Managers to make this comparison because of time constraints on the survey instrument.

Figure 2.4: Tasks of Supervisors: Perceived Ability by Gender

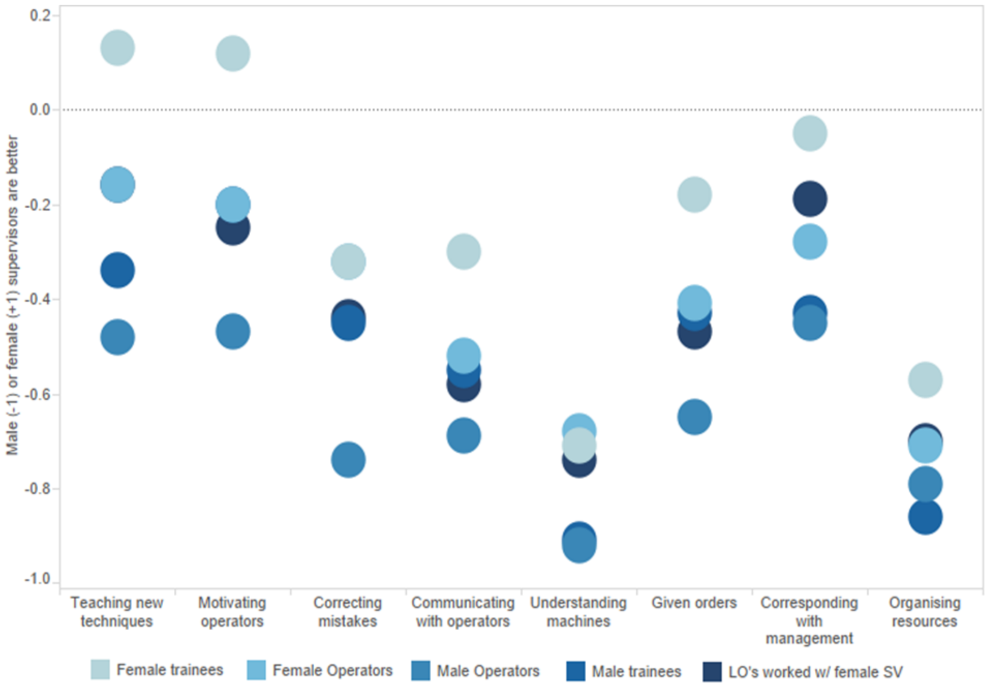


Notes: Workers on various levels in 26 factories were asked for each of the eight main supervisor tasks, whether they perceive female or male supervisor as more capable. Answers were aggregated on the task and designation of respondent level, with answers being coded as -1 for "males are more capable", 0 for "both are equally capable", and 1 for "females are more capable".

worse than the typical supervisor on each of the eight characteristics, while males rate themselves better at motivating workers and giving orders. The average gap for males is only 0.09, while for females it is 0.45. Across skills, the females' self-assessments largely match the pattern of the gender perceptions more generally. The correlation between the gaps the female trainees perceive in themselves and the gaps that operators perceive in female supervisors is 0.68.

We aggregate the ratings of males and females on all eight skills to create a single variable indicating each respondent's beliefs about the relative skills of males and females. For the aggregation, we assign a value of 1 to "females are better", 0 to "males are better" and 0.5 to the indifferent response. The first column of Table 2.2 shows how the average deficit for females across the eight tasks is affected by the gender of the operator and past experience working with female supervisors. Consistent with the data in Figure 2.5, we find that female operators have slightly

Figure 2.5: Tasks of Supervisors: Perceived Ability by Gender, Extended



Notes: Workers on various levels, of different gender, and with varying experience of working under female supervisors in 26 factories were asked for each of the eight main supervisor tasks, whether they perceive female or male supervisor as more capable. Answers were aggregated on the task and group of respondent level, with answers being coded as -1 for "males are more capable", 0 for "both are equally capable", and 1 for "females are more capable".

higher opinions of female supervisors, being about 12 percent more likely to choose "female is better" over "male is better". Previous reported experience working for a female supervisor does not change the perceived skill level of females and males. In the second column, we spilt the experience effect by the gender of the operator. There is no effect for female operators, while there is a small effect for male operators (p-value 0.101).

We also asked operators whether they prefer to work for a female or male supervisor. Similar to the coding for skills, we code the responses as 1 for "prefer female", 0 for "prefer male" and 0.5 for indifferent. As a group, the operators say they prefer to work for male supervisors by a margin of about two to one. However, female operators are 17 percent more likely to say they prefer females, and those

Table 2.2: Attitudes toward female SVs: Baseline data

	(1)	(2)	(3)	(4)
	Females better than males, all 8 tasks		Prefer female SV to male SV	
Operator is female	0.122***	0.142***	0.175***	0.208***
	(0.017)	(0.022)	(0.043)	(0.046)
Experience working for female SV	0.011		0.122***	
	(0.017)		(0.039)	
Experience * female operator		-0.002		0.099*
		(0.019)		(0.051)
Experience * male operator		0.047		0.182***
		(0.028)		(0.057)
Obs	428	428	426	426
R-squared	0.19	0.19	0.18	0.18
Factory FE	YES	YES	YES	YES
Mean	0.25	0.25	0.32	0.32

Notes: Standard errors clustered at the production line level; regressions include age, education and marital status of the respondent. Statistical significance of differences in comparisons: *** p<0.01, ** p<0.05, * p<0.1.

with previous experience working for female supervisors are 12 percent more likely to say they prefer working for a female supervisor (Table 2.2, column 3). Again there appears to be, if anything, a somewhat stronger effect for male operators (column 4) – though as with the skills assessment, the gap between female and male operators is not statistically significant. Among the 140 female operators reporting experience working for a female supervisor, 40 percent say they prefer to work for males, 30 percent for females and 30 percent are indifferent. Among males with no experience working for females, the percentages are 81, 16, and 3.

In sum, the skills assessment provides little evidence that perceptions are influenced by experience. However, when asked to express a preference to work for male or female supervisors, previous experience working for women does appear to matter, especially for male operators.

2.4 Do measured skills match the perceptions?

The surveys indicate that female supervisors are viewed as less skilled than male supervisors in each of eight supervisory tasks. The female trainees see similar weaknesses in themselves. Do these perceptions match reality? We conducted an extensive skill assessment of the female and male trainees selected by the participating factories during their first day at the training centre. We administered tests of numeracy, literacy, and non-verbal reasoning. We also directly assessed technical skills and knowledge of machines, and conducted teaching, communication, and leadership exercises. The data from this assessment provide evidence on several dimensions of the actual skills gaps between females and males selected by factories as having supervisory potential. We use these data for two purposes. The first is to assess the extent to which perceptions match reality at the baseline. The second is to measure the effects of training and the trial period working as an assistant supervisor on the trainees' skills. For the latter purpose, we repeat some of the exercises at the end of training and after the factory trial period.

2.4.1 Baseline measures: Do the skills gaps match the perceptions?

The most direct and extensive comparison we can make between perceptions and reality is on the question of technical and machine knowledge. The assessment asked the trainees to name different parts of sewing machines, and to tell us which type of machine (e.g., flat lock, single needle, etc.) would be used for different sewing processes. We showed the trainees garments of the type they typically produce with faults in them, and asked them to identify what machine problem (e.g., loose thread tension) is the most likely cause of the fault. We showed the trainees pictures of production lines and asked them to identify issues where worker safety was being compromised. In all, the diagnostic included 86 questions. Male trainees answered 65 percent of the questions correctly, while female trainees a statistically indistinguishable 64 percent correctly.

We conducted a very similar exercise after training and then again after the trainees completed the trial in the factory. The first column of Table 2.3 shows results of factory fixed effect regressions using all three rounds of the assessment. For now, we focus on the top line of the table, which shows the difference between females and males on the baseline assessment. We see that females on average score one point lower on the 86-point scale, a difference which is highly insignificant. In other words, even though close to 90 percent of survey respondents say that male supervisors have more technical knowledge than female supervisors, we find no statistical difference between the female and male trainees selected by the factories.

We also conducted exercises to measure teaching, communication and leadership. In the teaching exercise, we divided the trainees into groups of four to six. We assigned each trainee the role of teacher in one round of the exercise, with the others being students. The teacher was given an abstract figure, which might be for example several triangles and circles with some coloured in. The teacher's task was to instruct the students to reproduce the figure using only verbal instructions. She could not show the figure to the students or use her hands. There are two types of outcome measures. The simplest is the number of drawings that were correct. The first row of column 2 on Table 2.3 shows that males obtain a slightly higher percentage of correct drawings, with the gap being marginally insignificant with a p-value of 0.10.

The second outcome from the teaching assessment comes from observations recorded by two enumerators observing the exercise. For example, the enumerators recorded whether the instruction was given at an appropriate pace, and the number of times the teacher explained the task in more than one way. We take six such observations and construct standardized measures for each assessment round. We then sum the six standardized indicator variables to create an index of "soft teaching skills". Column 3 on Table 2.3 shows a factory fixed-effect regression with this index as the dependent variable. We see no significant difference between males and females in baseline teaching techniques, though the standard errors are larger than we might like. We also note that the soft skills measure is not significantly associated with the harder outcome – the percentage of correct drawings – though the measured effect is positive (p=0.22).

We create similar 'soft' measures for the communication exercise and the leadership exercise. In the communications exercise, the trainees were asked to give a short speech on a topic related to rules in the factory, such as: "Describe to a new operator all the things that you need to do when your machine breaks". During the speech, the trainee was interrupted with questions on two occasions. (For example, "What should I do if I think I can fix the machine myself?"). Two enumerators recorded judgements on whether the trainees spoke clearly, at a reasonable pace, whether she had confidence, etc. The top row of column 4 in Table 2.3 shows that

Table 2.3: Training and Trial Effects by Gender

	(1) Aptitude score	(2) Percentage drawings correct	(3) Drawing, “soft” score	(4) Communic. “soft” score	(5) Leaders. “soft” score	(6) Self confidence
Female trainee,	-1.19	-0.09	0.22	-0.27	-0.60	-0.47*
baseline	(0.90)	(0.06)	(0.56)	(0.72)	(0.53)	(0.24)
Female trainee,	-0.69	0.13**	0.22	-0.70	-0.80	0.24
after training	(0.86)	(0.06)	(0.60)	(0.81)	(0.67)	(0.25)
Male trainee,	0.20	0.18***	0.10	0.21	0.09	0.50**
after training	(0.80)	(0.06)	(0.64)	(0.79)	(0.79)	(0.25)
Female trainee,	-4.93***	0.09	-0.43	-0.01	-1.09	0.32
after factory trial	(1.82)	(0.08)	(0.58)	(0.85)	(0.69)	(0.24)
Male trainee,	-0.37	0.05	0.24	-0.75	-0.07	0.58**
after factory trial	(1.25)	(0.07)	(0.69)	(0.96)	(0.76)	(0.24)
Baseline mean, male trainees	55.4	0.33	-0.15	0.16	0.35	-0.33
Fem.vs.Male., after train. (p)	0.28	0.50	0.86	0.30	0.30	0.32
Fem.vs.Male, after trial (p)	0.02	0.66	0.32	0.48	0.20	0.30
Observations	470	420	421	423	329	470
R-squared	0.191	0.125	0.098	0.068	0.125	0.133
Factory FE	YES	YES	YES	YES	YES	YES

Notes: Unbalanced panel of trainees measured on first day of training, last day of training, and at the end of the factory trial period. Statistical significance of differences in comparisons: *** p<0.01, ** p<0.05, * p<0.1.

female trainees perform insignificantly worse by these measures. Finally, in the leadership exercise we asked the group to create a production hierarchy, and then asked them to produce some 'products' using Legos. The precise hierarchy depended on the size of the group, but we measure whether there are differences across the genders in the probability of being appointed a management role, and in soft measures reflecting the extent to which the individual participated actively in the discussion. We find that males are significantly more likely to be appointed to management (75 percent vs. 32 percent, $p<0.001$), but (see the top row of column 5), we see that women score insignificantly lower by these measures.

The teaching, communication and leadership exercises were intended to measure important aspects of confidence and preparedness to lead a production line. We also asked the trainees questions which yield a self-assessed measure of confidence. We first asked them to rate the average supervisor in their factory on a scale of 1-10. We then asked them to rate themselves as a supervisor, two months after beginning the job. Here we find that, at baseline, the male trainees express more confidence in their ability. In the raw data, they rate themselves 0.33 points lower, while the female trainees rate themselves 0.79 points lower. The top row of column 6 in Table 2.3 shows a similar deficit for women of 0.47 points, controlling for factory fixed effects. Thus, while both the technical assessment and the leadership exercises show no significant differences between the female and male trainees, we do see differences in their self-reported confidence levels.

2.4.2 Training and Trialing effects

We repeated the teaching, communication and leadership assessments at the end of the six-week training period and again at the end of the factory trial period. For the latter assessment, the trainees returned to the training centre for a review day during which we conducted these assessments as well. Rows 2 and 4 of Table 2.3 show the various post-training measures, all measured relative to baseline. At the bottom of the table, we show the p-value for tests of equivalence of female and male trainees at each point in time. The table is an unbalanced panel, as there were several nominated trainees who either failed the literacy / numeracy exam or dropped out for other reasons, and there was further attrition before the post-trial review day. However, results from regressions using the balanced panel are very similar, suggesting that the patterns we observe are not driven by selection. Looking first at the scores on the assessment of technical knowledge, we see slight improvements in both males and females after the training, but no change in the relative performance across gender. After the factory trial (rows 3 and 5), however, we see the performance of females appears to deteriorate somewhat. Indeed, comparing female and male trainees, the post-trial technical assessment is the only measure showing a significant difference by gender. The other outcome worth noting is that confidence of both males and females increases after the training, with the measured magnitude of the increase for females being slightly larger, but not significantly so.

2.4.3 Management simulation exercise

In the first phase of the project, we conducted a management simulation exercise which we believe illuminates some of the issues facing female trainees. The exercise was conducted during a follow-up survey around four months after the completion of training. The simulation involved the trainees and eight randomly selected operators. The operators were placed into four teams of two each and played two "production" games, one involving Legos and one involving buttons. We randomized the order in which the games were played at the factory level. Each team was assigned a leader whose job was to explain the particular exercise and manage the operators as they performed their tasks. For the results we present here, the team leader was either a female or male trainee.[15] Each pair of operators played the production game twice, once with Legos and once with buttons. Each team leader played only one session – either Legos or button – so there were eight team leaders in each factory, and each pair of workers played with two different team leaders.

For each of the Lego and button exercises, the teams played five separate sessions. The first was a simple sorting exercise in each case, sorting either buttons or Legos by colour. For Legos, the second, third and fourth sessions involved constructing chains of Legos with a particular color pattern – blue, yellow, green, blue, yellow, green, etc. The three games were differentiated by their payoffs: the first summed the length of the chains produced by the two operators, the second paid based on the length of the longest chain produced by either worker, and the third paid based on the shortest chain produced by either operator. The team leaders were given incentive payments according to the payoff function.

[15] The full exercise also involved team leaders who were operators from the control group, the most recently promoted supervisor who was not a trainee, another supervisor from the same production line as one of the trainees selected at random (a "matched" supervisor), or another supervisor selected at random (a "random" supervisor).

Here we assess the performance of teams led by female trainees with that of teams led by male trainees measured by the payoffs. We combine each of the five individual games into a single regression by standardizing the payoffs on the game-round level. We then run regressions with the standardized payoffs on the left-hand side and a set of controls for characteristics of the team leader on the right hand side. We focus the discussion here on the subset of games where trainees are team leaders, comparing the performance of female and male team leaders.

Each pair of operators plays the set of five games twice, with one team leader in the first session and a different team leader in the second session. The order of the games (Lego - buttons, or buttons - Lego) is random, and within a round the assignment of team leaders to operator pairs was random. But the assignment of team leaders to session 1 or round 2 depended on the (non-random) order in which leaders were provided by the factories. Logistical complexities working in the factory prevented us from randomizing the session in which any team leader participated. In particular, because factories anticipated that we wanted to talk with trainees, the trainees were more likely to be assigned to the first session, and the existing supervisors and control operators were more likely to be assigned to the second. This matters, because even controlling for the team leader and game types (Lego vs. buttons), operators were significantly more productive during the second session. This is logical because we expect some learning by the operators from the first to the second session – even though they play different games in each session. We control for the session order effects in regressions.

Table 2.4 shows how the standardized payoffs vary with the gender of the team leader in the sample of games involving female and male trainees. The specification in column 1 includes controls for factory, session (first or second) and game fixed effects. We find that teams led by female trainees have payoffs which are 0.29 standard deviations higher than teams led by male trainees, a difference which is highly significant. In other words, female trainees appear to be more effective as team leaders than male trainees. Column 2 adds team leader demographics – age, education, industry experience and factory tenure – and Column 3 adds operator team fixed effects. Note that the third regression then isolates the cases where a single team was led by both a male and a female trainee. Only 19 teams had this pair of team leaders, so while the table shows a sample size of 600, the effective size is much smaller. Nevertheless, the additional production by female-led teams is statistically the same, increasing to 0.42 standard deviations.

In columns 4 and 5 we examine whether trainees who before the survey visit had been tried out as supervisors or promoted to supervisor perform better than those not tried out or promoted. We find that those promoted to supervisor perform significantly better than those not promoted. Since promotion is not random, we are unable to say whether this is due entirely to selection – more able trainees are promoted, while less able ones are not – or whether the experience as a supervisor also makes the individual more effective as a leader. But the result does provide some validation for the exercise itself, showing that those with more ability or experience perform significantly better in the game.

Finally, in column 6 we explore whether the gender composition of the operator team interacts with the gender of the team leader. We compare the performance of mixed or all-male teams with those of all-female teams.[16] The superior performance of the female-led teams is significant only for the all-female team. When both operators are women, their output is 0.41 standard deviations higher with a female leader than with a male leader. But female leaders obtain no higher production from mixed team than male leaders do.

In sum, then, the female trainees were significantly more effective in generating payoffs than were the male trainees. Trainees who were promoted before the time of the first follow-up survey perform significantly better than those not promoted. And female trainees perform best when they are matched with a pair of female operators, and perform no better than male trainees when they lead mixed-gender or all-male teams.

There are two further outcomes from the games. The first involves the strategy choices of the team leaders. Recall that the payoffs changed from one game to the next. In the second round, payoffs were for the sum of the output of the two operators, but the third (fourth) game, we paid on the maximum (minimum) output of either worker. After the second game, we asked each team leader which of the two operators was better at the game. We then recorded whether the team leader focused attention on the stronger operator in game 3 and on the weaker operator in game 4, as the performances of these operators are expected to determine the payoffs in these games. The fifth game involved a complex figure that was most efficiently made in a "line", with each operator specializing on one component. We

[16]Since only 20 percent of operators are male, all-male teams are very rare.

Table 2.4: Management Simulation Exercises: Female vs. Male Trainees

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Standardized Pay–Off in Games						Correct Strategy
Female Team Leader	0.290***	0.255**	0.420**	0.305**	0.332***	0.412***	-0.371**
	(0.109)	(0.122)	(0.190)	(0.131)	(0.124)	(0.521)	(0.167)
Mixed gender / male team						-0.092	
						(0.725)	
Mixed * Female TL						-0.433*	
						(0.240)	
Tried as Line Supervisor				0.329			
				(0.212)			
Promoted to Line Supervisor					0.508**		
					(0.206)		
Team Fixed Effects	no	no	yes	no	no	no	no
Game Fixed Effects	yes	yes	yes	yes	yes	yes	yes
Team Leader Demogr.	no	yes	yes	yes	yes	yes	yes
Number of Observations	676	612	612	612	608	612	612

Notes: The dependenet variable is the standardized payoff from the game. Standard errors are clustered at the game level: *** p<0.01, ** p<0.05, * p<0.1.

.

record whether the team leader organized production in that manner. We then sum the number of times the team leader adopted the "correct" strategy in each of these three games. Column 7 regresses this sum on the gender of the team leader, demographics of the team leader and the operators, and factory fixed effects. We find that the male leaders adopted the correct strategy significantly more often, in spite of the female leaders inducing higher output.

Finally, after the second session, the operators on the production team were asked to compare the management style of the two team leaders they worked with. They were asked whether the first or second team leader they worked with was better at explaining the game, better at answering questions, better at motivating them, always pressuring them, and so forth. Focusing on the responses of the 19 teams led by both a female and a male trainee, we find that operators are more likely to say that the male trainees were better at answering question, at motivating, and at encouraging. Female trainees were selected more often only as "always pressuring". The last two outcomes, on strategy and operator opinions, are interesting in the light of the superior performance of female trainees.

2.5 How are operator perceptions changed by experience?

The skills diagnostics indicate that the female trainees have only a very small and statistically insignificant gap in technical skills. On the other hand, there are more significant gaps in self confidence and in the outcomes of the teaching and leadership exercises. The training closes these gaps. But the important outcomes are not the training centre diagnostics, but the outcomes on the production floor. We examine these using both surveys of operators working for the trainees and using administrative data on productivity of the lines where the trainees are assigned (ITT) or work (OLS).

We conducted a first follow-up survey in the factory just after the end of the initial trial period. During the six- to eight weeks between the end of the training and this first follow-up survey, trainees were meant to be working as assistant line supervisors, together with their mentor. Compliance with this agreement was very high. Of the 135 operators completing training 129 were trialed as an assistant supervisor. Four of the six not trialed (three females and one male) left the factory before the trial started. Recall that we told factories which specific trainees to place on which lines. However, arguably we should only be concerned that they placed a female (male) trainee on a line assigned to a female (male). In many cases, the factory did not comply with the assignment at the individual level, but did comply with the assignment at the gender level – that is, they switched two females or two males. The trial was carried out on a different line in 34 percent of the cases. But in the 77 cases where trainees were trialed on the lines designated for trials, there was compliance in 71 cases.

At the follow-up conducted at the end of the trial period, we surveyed randomly selected operators working on the lines where the trainees were assigned (ITT lines) and on the lines where the trainees were actually working. Recall that for each factory, the training was conducted in two rounds approximately two months apart. The first follow-up survey was also conducted twice in each factory, at the end of the factory trial for each training round. The second follow-up survey, however, was conducted on both training lines at the same time. This has two implications. First, this implies that the time gap between the end of the trial and the second follow-up survey was about two months longer for the trainees in the first training round than for those in the second round. Second, this meant that we were surveying twice as many lines on the day of the second follow-up. As a consequence of this, we did not survey all of the ITT lines at second follow-up. Therefore, we report both ITT and OLS regressions for the first follow-up survey data, but only OLS regressions for the second follow-up survey.

At each follow-up survey, we selected three operators at random from each of the surveyed lines. We focus on two outcomes. First, we asked the operators to rank on a scale of 1-10 both a typical supervisor in the factory and the trainee on their line based on their knowledge of her/him. We regress the ranking of the trainee on an indicator for his or her gender and for the gender of the surveyed operator, controlling for the ranking of the typical supervisor by the operator. Second, we asked the operators whether they prefer to work for a female or male supervisor, and as before code the responses as 1 for “prefer female”, 0 for “prefer male”, and 0.5 for “indifferent”. For the first of these outcomes, we are interested in the ranking of female trainees relative to male trainees, and for the second, we are interested in whether exposure to a female trainee affects the preference for supervisors.

The first three columns of Table 2.5 below show the ITT regressions for the relative ranking (columns 1 and 2) and the preference for female supervisors (column 3). We find that the female trainees are rated almost a point – about 0.4 standard deviations – lower than the male trainees. In column 2, we allow the relative ranking to differ for female and male operators. We find, if anything, males rate the females more harshly, though the difference is not statistically significant (p=0.26). In column 3, we see that exposure to female trainees has the effect of making male operators significantly less opposed to working with female supervisors. While female operators are more inclined than male operators to say they prefer to work for female supervisors, their opinion is not influenced by exposure to the female trainees.

Columns 4-7 of Table 2.5 repeat the same regressions using the actual placement of the trainees. We find almost identical effects in the ranking regressions (columns 4 and 5), but slightly weaker effects in the preference regressions (column 7). Finally, columns 8-10 show the results of OLS regressions using the second follow-up survey data. Because we use the sample of trainees working as assistant supervisors or full supervisors at the time of the second follow-up, in column 6 we show the first follow-up results using the sample of trainees working as supervisors at the time of the second follow-up. We see that the results for male operators are very similar to those in the full sample (compare column 6 with column 5), though the smaller sample yields higher standard errors and an insignificant effect. The results for female operators appear slightly different and less negative, for the sample of trainees that continue to work as supervisors at the second follow-up. This indicates that the weaker female trainees may be those who do not continue as supervisors.

In the second follow-up survey, the deficit for female trainees is erased completely (See columns 8 and 9 of Table 2.5). Female trainees are rated as equal to male trainees, by both female and male operators. Moreover, operators of either genders who are exposed to the female trainees express higher preferences for working with female supervisors. Note as well that the trainees as a whole are now rated as slightly better than the typical supervisor in the factory. The improvement in the relative ranking of the trainees is consistent with statements by production managers that new supervisors require four to six months of experience to reach their full potential.

Table 2.5: Comparison Female and Male Trainees by Worker, Follow-Up Surveys

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	MIDLINE							FOLLOW-UP 2		
	ITT			OLS				OLS		
	Rank (1-10) for Trainee		Prefer Male SV	Rank (1-10) for Trainee			Prefer Male SV	Rank (1-10) for Trainee		Prefer Male SV
				— Full Sample —		SV at FU				
Female trainee on line	-0.90***			-0.91***				0.06		
	(0.25)			(0.22)				(0.29)		
Female trainee on line * Female operator		-0.65**	-0.01		-0.79***	-0.38	0.02		0.04	0.16***
		(0.31)	(0.06)		(0.26)	(0.30)	(0.05)		(0.31)	(0.05)
Female trainee on line * Male operator		-1.39***	0.18***		-1.19**	-1.04	0.12**		0.13	0.17*
		(0.52)	(0.06)		(0.46)	(0.65)	(0.06)		(0.57)	(0.10)
Operator is female	0.13	-0.25	0.24***	0.15	-0.03	-0.20	0.16***	0.48*	0.51	0.16**
	(0.32)	(0.52)	(0.13)	(0.27)	(0.39)	(0.50)	(0.05)	(0.28)	(0.34)	(0.07)
Ability (1st PC)	0.00	-0.01	0.003	0.14	0.14	0.23**	-0.01	0.20*	0.20*	0.01
	(0.10)	(0.10)	(0.014)	(0.11)	(0.11)	(0.11)	(0.01)	(0.11)	(0.11)	(0.014)
Rank typical supervisor	0.41***	0.41***		0.35***	0.35***	0.47***		0.35***	0.35***	
	(0.09)	(0.09)		(0.07)	(0.07)	(0.08)		(0.08)	(0.08)	
Observations	238	238	357	295	295	203	410	206	206	215
R-squared	0.221	0.226	0.116	0.232	0.233	0.314	0.093	0.279	0.279	0.201
Mean male ops., preference			0.22				0.22			0.25
Avg ranking Typical SV				7.41	7.41	7.41		7.58	7.58	
Avg ranking of Trainees				6.86	6.86	6.86		7.82	7.82	
Line FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Ability is the first principal component of years of schooling and scores on the numeracy, literacy, non-verbal reasoning and technical aptitude test.

2.6 Trainee Performance measured by Production Data

The literature measuring the effects of job training programs has typically relied on outcome measures such as employment or earnings of trainees.[17] This is reasonable if wages equal the value of the marginal product of labour. In our context, we believe that approach has drawbacks. First, the factories typically have very specific wages for each worker grade. Many or most of these are determined by minimum wage levels, which are set nationally at the worker grade level. Thus, wages may not reflect marginal products. Second, factories will attempt to make promotion decisions based on their beliefs about actual productivity of the workers.

With this in mind, we have attempted to gather very detailed production data for each of the factories. For the second phase of the project, we have daily, line-level data for 12 or 13 months, typically starting two months prior to the beginning of training and extending seven to nine months after the end of the training (see Appendix B for a more detailed description of the data and its collection process). There are three outcomes of interest: productivity, quality defects, and absenteeism. By focusing on sewing, we are able to capture a measure of output which is very close the pure quantity measure. A trained industrial engineer can take any garment and estimate the number of minutes a fully-efficient worker will take to produce the garment. These calculations come from summing the required time for each stitch to make the garment. The times come from a combination of international databases and in-factory time-and-motion studies. By multiplying these 'standard minute values' – SMVs (or standard allowable minutes – SAMs) by the number of units of a given garment which are produced during the day, we obtain a measure of output – output minutes – which is highly comparable across products. For example, a line producing 1,000 shirts with an SMV of 15 minutes has production of 15,000 output minutes. For productivity, we divide the output minutes by input minutes – the sum of minutes worked by operators and helpers on the line over the same time period[18] – to obtain the industry standard measure of efficiency. This is essentially a measure of Q/L:

[17] Much of this literature focuses on programs aimed at individuals who are out of work. See, for example Card et al 2011; Attanasio et al 2011.

[18] Helpers are entry level workers who are allocated to lines but do not yet operate a sewing machine on their own. They typically represent the lowest wage grade in the factory and do auxiliary tasks on the lines, such as cutting thread or moving garments from one operator to the next. We could improve the input minutes measure by a step if we had the wage bill for the whole line. However, the industry typically uses three different wage grades for operators, and we most often know only the total number of operators, not the number by grade.

$$Output * SMV/[(Operators + Helpers) * hours * 60] \tag{2.1}$$

The average efficiency in the sample we are currently using is 53 per cent, which is higher than the 38-40 per cent that those in the industry typically quote.[19]

A second measure of interest is the number of quality defects. Factories typically report both the number or percentage of garments that require some re-work and the number or percentage that must be rejected. Reject rates are typically very low, averaging less than 0.5 percent in our sample. Rework rates are much higher, averaging around 7 per cent (with a median of almost 5 per cent). Because the re-work time is included in the measure of "input minutes", the efficiency measure incorporates improvements in quality.

We construct a panel at the line level, with dummy variables indicating the presence of a trainee working on the line either as an assistant supervisor or a full supervisor. We begin with an ITT specification, using the gendered assignment of a trainee on the line during the trail period, and then assuming this initial assignment predicts the line on which the trainee will be promoted.

$$y_{gfld} = \alpha_l + \beta_{fd} + \sum_{g\in\{0,1\}} \gamma_g Trial_{fld} + \sum_{g\in\{0,1\}} \delta_g Post\_Trial_{fld} + \epsilon_{gfld} \tag{2.2}$$

where $g \in \{0, 1\}$ represents male or female trainees, f is factory, l line, d the week of production, and y the outcome of interest. $TRIAL$ reflects the assignment of the line to a female /male trainee during the trial weeks and $POST\_TRIAL$ the assignment of the line to a female/male trainee during the period after the trial.

We also present OLS results on the actual placement and roles of trainees. These may suffer from both the endogenous placement of trainees and the endogenous decisions to promote. As with the ITT regressions, we include both line and factory/week fixed effects, which mitigates to some degree the issue of endogenous placement. However, some of the trainees leave the factory and some return to being operators after the trial. Since these outcomes are more frequent for females than

[19]The higher efficiency in our sample may come from having a more efficient sample of factories. However, the data across factories are not always comparable because the international SMV values are often adjusted upwards by factories to account for some expected level of inefficiency. We are currently working to ensure the data are comparable across factories, but we include factory fixed effects in all of the regressions using production data, which will absorb systematic measurement differences across factories.

for males, we should clearly be concerned with the endogenous promotion decisions in interpreting the OLS regressions. We nevertheless think that the OLS results are potentially interesting in spite of these selection issues, because promotion of almost any females represents a change relative to what would have happened in the absence of the experiment.

The first three columns of Table 2.6 report the ITT regressions for efficiency, absenteeism and defect rates. The samples for each of the regressions vary somewhat because data on some measures are not available in some factories.[20] The cleanest results relate to efficiency. Compared to lines without trainees, we see that lines where male trainees were assigned are about 2.3 percentage points – roughly 5 percent – more efficient during the trial period. During the trial, the trainees represent extra supervisory labour on the line. Hence, even though they are least experienced at this point, it is perhaps not surprising that they have a positive effect on efficiency. There is no increase in efficiency during the trial period on the lines assigned a female trainee, suggesting that even though the female trainees are additional supervisory labour, they are not effective in increasing efficiency. However, the situation changes during the post-trail period. Those trainees remaining as supervisors may either be classed as Assistant Supervisors or as full Line Supervisors during this period. In the latter case, and perhaps even in the former, they are replacing an existing line supervisor, and hence no longer represent incremental supervision. During this period, the female trainees catch up to the males. We see that both female and male trainees have very similar effects on efficiency, with positive coefficients which are economically important but statistically insignificant at conventional levels.

Columns 4 through 7 present OLS results based on actual assignment. We use actual assignment because the initial line assignment was agreed to only for the trial period. We did not necessarily expect the factories to promote the trainees to the same lines. The patterns are very similar to the ITT regressions, though the coefficients are generally of slightly larger magnitude. The regression in column 4 shows that females perform significantly worse during the trial period, perform equally well as males when both are assistant supervisors, and perform insignificantly better than male trainees when both have been promoted to full supervisor. In column 5, we limit the sample to observations from days when the trainee was

[20]The sample size drops by about 75 percent when we use lines for which all three variables are not missing.

Table 2.6: Productivity of Trainees on the Line

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MODEL	— — — —	ITT	— — — —	— — — —	Actual Assignments	— — — —	
VARIABLES	Efficiency	Absenteeism Rate	Defect Rate	Efficiency	Efficiency ITT lines	Absenteeism Rate	Defect Rate
Trial period, Female trainee	-0.0093	-0.0010	-0.0020	-0.0237	-0.0174	0.0007	-0.0013
	(0.0118)	(0.0054)	(0.0037)	(0.0165)	(0.0180)	(0.0035)	(0.0029)
Trial period, Male trainee	0.0236**	0.0050	-0.0042	0.0190	0.0112	0.0070	-0.0043
	(0.0108)	(0.0046)	(0.0036)	(0.0146)	(0.0154)	(0.0046)	(0.0032)
Post-trial, Female trainee	0.0176	-0.0004	-0.0055*				
	(0.0150)	(0.0039)	(0.0032)				
Post-trial, Male trainee	0.0196	-0.0025	-0.0013				
	(0.0128)	(0.0030)	(0.0035)				
Assistant SV, Female Trainee				0.0250	0.0220	-0.0049	0.0043
				(0.0177)	(0.0170)	(0.0083)	(0.0042)
Assistant SV, Male Trainee				0.0240**	0.0302	0.0082**	-0.0057
				(0.0121)	(0.0238)	(0.0036)	(0.0049)
Line SV, Female trainee				0.0101	0.0369	0.0039	-0.0036
				(0.0153)	(0.0244)	(0.0043)	(0.0037)
Line SV, Male trainee				0.0015	-0.0060	-0.0045	0.0006
				(0.0113)	(0.0147)	(0.0043)	(0.0035)
Observations	78241	70814	70814	78241	18888	70814	87952
R-squared	0.54	0.39	0.39	0.53	0.56	0.39	0.59
Number of factory Lines	419	264	431	495	495	495	495
Line FE	YES	YES	YES	YES	YES	YES	YES
Factory/week FE	YES	YES	YES	YES	YES	YES	YES
TEST: Fem.vs.Male trainee, Trial	0.027	0.375	0.663	0.045	0.179	0.293	0.475
TEST: Fem.vs.Male trainee, Asst.SV	0.916	0.636	0.336	0.962	0.792	0.141	0.111
TEST: Fem.vs.Male trainee, Line SV				0.631	0.096	0.185	0.435
Nr. Lines assigned Male Trainee:	49	32	51				
Nr. Lines assigned Female Trainee:	43	28	44				

Note: Robust standard errors clustered on the line level; *** p<0.01, ** p<0.05, * p<0.1

working on one of the original ITT lines. The patterns are similar, though now the better performance of female trainees as full supervisors is marginally significant (p=.096). Columns 6 and 7 report results for absenteeism and defect rates, respectively. Again the patterns are similar to the ITT regressions except that underperformance of female trainees relative to male trainees on quality issues is almost significant when working as assistant supervisors (p=0.111; see bottom of table).

The efficiency results mirror the opinions of operators working on the lines. Female trainees start slower; they perform significantly worse than males during the trial period. However, they catch up in the months after the trial period. We see the same pattern in the ITT and OLS regressions. In the ITT regressions, the gain made by female trainees relative to male trainees is significant at the 0.10 level, while the gain in defect rates is marginally insignificant (p=0.125).[21] In the OLS data, we find significant relative performance gains between the trial and promotion to line supervisors for efficiency, and between the period working as an assistant supervisor and promotion to line supervisor for defect rates.

2.6.1 Do attitudes adjust?

Both the survey data and the production data suggest that the female trainees start more slowly than their male counterparts, but catch up three to five months after returning from training. The attitudes of operators with direct exposure to the female trainees adjust over this time. We might ask whether there is any evidence that the attitude adjustment is more general. That is, does the increase in female supervision in the factory have indirect effects on operator attitudes towards female supervisors? The data suggest there is no change to attitudes of other workers: The sum of the generic female / male rankings – coded, as before, 1/0/-1 – is -5.06, -4.97 and -5.19 for the male operators surveyed at baseline, first follow-up and second follow-up, respectively, and -3.33, -3.46 and -3.28 for female operators at the same surveys. None of these differences are statistically significant. Direct exposure to female trainees, on the other hand, has a significant effect on these rankings by the time of the second follow-up survey: Female (male) operators working on lines with female trainees have a cumulative ranking -2.86 (-4.71), compared with -3.71 (-5.28) for those on lines with a male trainee. The female operator gap is significant with

[21] This is the p-value comparing the gap between female performance in the post-trial period and female performance in the trial period with the same gap for males.

a p-value of 0.07. Generic attitudes show some evidence of movement with direct exposure, but there is no evidence of any broader effect in the factory.

2.7 Conclusion

This chapter set out to understand some of the reasons for the very low share of female production supervisors in the Bangladeshi garment industry, despite women representing 80% of the industry's workforce. We partnered with local training centres for mid-level managers, implementing a project which trains equal numbers of male and female workers to become sewing line supervisors and subsequently placed them as assistant supervisors on randomly selected lines at their factories for a trial period of two months. We accompanied the project with extensive surveys of workers, supervisors and managers at various levels in the factories, to gain additional insights into the factors which keep the number of women in supervisory positions in the industry that low.

Several key findings emerge from our analysis which provide novel insights into this question. First women are generally perceived to be less competent when it comes to solving technical and organizational problems. And while this perception is more pronounced among male workers, it is also prevalent among females. On the other hand, women are considered almost on par when it comes to leadership and communication skills. We compare these findings against results from extensive knowledge and skills tests of the workers nominated for the training, which reveals a stark contrast: Women have equal technical knowledge, but lag behind to some extend in leadership skills, and strongly when it comes to their confidence in their own supervisor skills.

This mismatch between perceived and actual weaknesses of females as supervisors could prevent factory managements to take effective measures to bring more women into supervisory roles. We show that the training sessions and the two months trials closed the gender gap in terms of confidence, and initial productivity difference of the trial lines to which female trainees were assigned, compared to the lines of male trainees, vanished after two months. Furthermore, workers directly exposed to female supervisors improved their rating of them. This indicates that an external intervention in which management puts trust into female supervisors over a few months of initial trial time can overcome the initial lack of confidence, expressed by both female candidates for supervisor positions themselves, and other workers in the factories, into their supervisor abilities. However, it could be precisely the documented misperceptions about relative supervisory ability of men and women which so far prevented more factory managements from taking such measures.

Chapter 3

Organizational Learning: Experimental Evidence from Bangladeshi Garment Factories

3.1 Introduction

Learning on the job has long been known as a key driver for productivity growth (Arrow [1962]; Lucas [1993]). Conceptually, learning on the job, especially within organizations, can be separated into learning by doing something oneself, and learning from co-workers. While the evidence on the first is extensive and goes well back into time (Wright [1936]; Benkard [2000]; Hendel and Spiegel [2014], see Thompson [2010] for a review), there is only limited evidence available on the second. Levitt et al. [2013] in a US car plant, and Thompson and Thornton [2001] in US ship yards show that productivity of workers not only increases with the amount of a certain car or ship model they produced themselves, but also with the amount of the product produced by others in the firm. However, beyond the mere documentation of the existence of learning from others in firms, little is known about, for example, under which conditions it works particularly well.

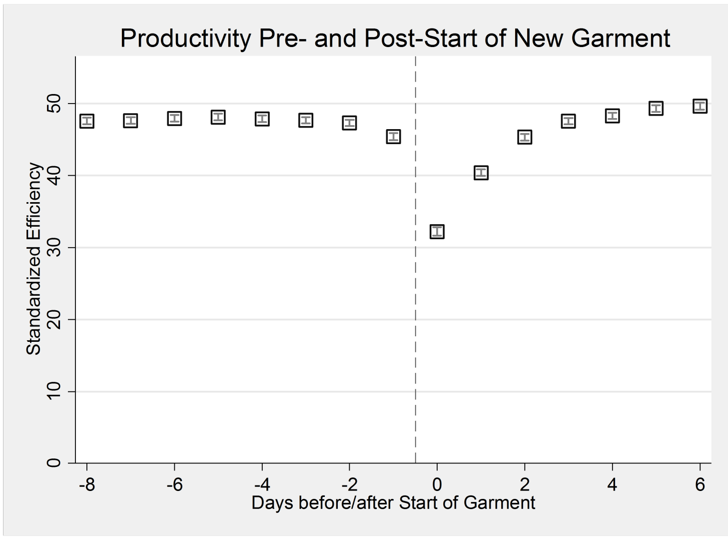
One reason for the paucity of evidence on learning from co-workers is that knowledge exchange between people is inherently difficult to observe. The above mentioned studies measure organizational learning through increases in productivity across production units and time, if other workers in the organization have already produced the product before. However, this leaves open the possibility that the productivity increases are not driven by knowledge transfers, but alternative forms of peer effects.[1] To clarify the mechanism behind the cross-unit productivity increases within firms, I collect data and run a randomized experiment in three Bangladeshi garment factories. In the experiment, random pairs of workers are induced by their superiors to brief each other when one of them starts producing a garment that the other one has previously produced. This communication intervention introduces exogenous variation in knowledge exchange across worker pairs, and I show that this intervention increases the productivity of the workers receiving the briefing. This indicates that increases in knowledge exchange among co-workers can increase productivity, above the levels achieved if workers would just learn from doing a certain task by oneself. Furthermore, there is some evidence that the effect of the intervention was stronger where the worker being briefed shared social ties with the worker who provided the briefing.

Bangladeshi garment factories are an ideal setting for studying organizational learning. The sewing departments of these factories, on which this paper focuses, are organized into parallel sewing lines of 20-80 workers, which are designed as independent production units on which the whole sewing process of a garment can be completed. The three factories in my sample have more than 200 sewing lines among them. Due to large order sizes and tight delivery deadlines, most garments are produced on more than one line. In these cases, the different lines typically start producing the garment on different days, as they finish previously allocated jobs. Thus, when a given sewing line starts producing a new garment, there may or may not be other lines in the factory with production experience on that garment. When there are such lines, they have gained potentially valuable production information which may or may not be shared with the sewing line starting production at a later time.

Sewing lines switch to new garments with relatively high frequency, on average every 10 days. Due to the fast-moving fashion industry and its seasonality, garments are technically differentiated, which is reflected by line productivity dropping by a third on average when lines switch to a new garment. Only four to five production days after the start of a new garment does line productivity reach its long run level again (Figure 3.1). Given these learning processes and fast turnover

[1] This problem also holds for most studies on social learning outside organizations, for example about new agricultural technologies (Foster and Rosenzweig [1995]; Munshi [2004]; Bandiera and Rasul [2006]; Conley and Udry [2010]), or about microfinance services (Banerjee et al. [2013]; Cai et al. [2015]), which use observed technology adoption as proxy for learning. These studies mostly rely on estimating structural models or on placebo tests to separate learning from other possible peer effects.

Figure 3.1: Sewing Line Productivity before and after Start of a New Garment.



Graph shows average sewing line productivity in the days before and after switching to a new garment. The vertical dashed line denotes the switch to the new garment, and Day 0 the first day of production of the new garment. The capped bars represent 95% confidence intervals.

rates of garments, the potential gain from knowledge spill-overs is large. Workers - or at least the supervisors of the lines who are held accountable for the productivity of their lines - should have a strong incentive to utilize this knowledge.

The randomized communication intervention was conducted among the line chiefs, the person responsible for the overall management of a line. For a period of four months on randomly selected sewing floors in the factories, line chiefs were instructed by the factory management to brief each other when one of them started producing a garment that the other one had previously produced on his or her line. The briefings were designed to last about 20 minutes, during which the line chief with experience should have shared the most important production problems which had to be overcome when producing the garment. I show that these briefings increased productivity of the sewing lines of the line chiefs who received the briefing during the first one to two days the garment was produced on the later line, before productivity reached its long run level again.

The randomized intervention covers only a fourth of the dataset of sewing line productivity that I collected. Using data from as long as two years before the experiment began, I first document that sewing lines are more productive on the first days they produce a new garment when that garment has already been produced on other sewing lines before. This effect is stronger if the garment has been produced on lines which are located spatially closer to the line starting the garment. However, the effects I document in the observational data could also be driven by other forms of peer effects, such as competition. For example, I show that sewing lines are similarly more productive on the first days they produce a new garment when they are the first line in the factory producing it, and other sewing lines start to produce the same garment on the same day as well. This effect is also stronger if these lines are located spatially closer. In these instances, no other workers have previous experience with the garment that could be shared. In principle, these contemporaneous effects could be driven by the selection of garments which are produced on multiple lines on the first day, for example those with especially close delivery dates. Nevertheless, this raises the question of whether the productivity increases observed when other lines have *prior* experience producing the same garment might also be generated by the same forces. By exogenously increasing the potential for knowledge exchange between workers, the randomized intervention therefore provides valuable evidence to what extent learning across co-workers contributes to these productivity increases.

Using survey-based information on the social network among line chiefs, I find some evidence that the effect of the briefings was stronger among line chiefs who shared social ties. Furthermore, the overall effects of the intervention are stronger when conditioning on line-garment type fixed effects. This could indicate that the intervention had a stronger effect when line chiefs started new garments of types that they would otherwise have struggled more with. This also fits with the overall effect of the intervention being driven by a reduction of the left tail of productivity; the intervention reduced the number of starts of garments on lines with extremely low efficiency on the first days.

An advantage of this setting is that the dataset I collected allows me to observe the production and knowledge diffusion process of more than 1,000 different garments over the same set of sewing lines, while earlier studies on organizational and social learning typically observe the diffusion process of a single product, or a small number of products. This allows me to control for time-invariant differences between sewing lines when comparing productivity. Furthermore, I can check whether the diffusion process of the garment is affected by observable product characteristics. Finally, the production data from these factories contains accurate productivity data for the sewing lines which has been standardized across different garments. Therefore, I do not need to base the comparison of output across heterogeneous products on assumptions which can be difficult to test (Foster et al. [2008]).

The positive effects of the intervention raise the question why similar management practices have not been implemented before at the factories. This question relates to the literature on management practices in large firms (Bloom et al. [2013]; Bandiera et al. [2015]), which finds that firms, especially in developing countries, fail to adopt practices which should be universally beneficial. A post-experiment survey conducted with the head production engineers of the three factories, who supervised the implementation of the experiment, revealed that the costs of the experiment in terms of necessary labour input time was negligible compared to the productivity gains. Instead, two of the three engineers reported they had never thought of conducting such an experiment before, while the third had but did not expect it to yield enough benefits. Finally, two out of the three factories continued with the communication practice after the end of the experiment, with the third citing resistance from line chiefs as reason for the discontinuation. This could point to a conflict between who incurs the gains of the intervention (the firm, through higher productivity) and who bears most of the costs (the line chiefs, in terms of the opportunity costs of the time it takes to provide the briefing), which resembles findings from Atkin et al. [2015] on non-adoption of new technologies in Pakistani factories. Furthermore, the fact that the intervention seemed to have a stronger effect when provided by socially connected workers points towards a broader aspect, that many management practices rely on non-verifiable cooperation of workers for their effectiveness, such as efficient communication of relevant production information. However, workers might not have incentives for such cooperation for all kind of reasons, rendering such seemingly universally beneficial management techniques less effective in reality.

Apart from the literature on organizational learning mentioned above, this paper also relates to small but growing literature on experiments within firms. Atkin et al. [2015] vary the pay-schemes for workers in Pakistani soccer ball factories, and show that this variation affects whether employees report truthfully about the benefits of new technologies in the production process. Bloom et al. [2013] randomly selected textile factories in India who received in-depth management consulting, and showed large effects on productivity of the firms. They cite low competitive pressures as reason for the previous non-adoption of these practices by the factories, as well as a lack of trust to managers from outside the owner family, who could introduce better management practices. Bandiera et al. [2013] introduce rank incentives and a tournament for production teams at a soft fruit producer and demonstrate that while the first measure decreases productivity, with the worst teams becoming even less productive, the latter increases overall productivity, with the effect being driven by the best teams becoming more productive. Bandiera et al. [2005] compare worker productivity under piece rate pay and relative payment at the same firm. They show that relative pay reduced worker productivity, but only in cases in which workers can monitor each other's effort. Bandiera et al. [2011] provide an overview over this literature.

By using detailed data on worker and productivity at the sub-firm level, this paper also connects to a broader literature on the interplay between management, worker characteristics and productivity. Amodio and Carrasco [2015] exploit exogenous variation in worker productivity in a setting with quasi-team incentives, and show free-rider effects among co-workers, with the effect being ameliorated either by social ties between workers or the introduction of piece rates per worker. Hjort [2014] is a case study of a Kenyan flower packaging factory showing that ethnically diverse work teams have lower productivity than ethnically homogeneous teams. The negative effect of heterogeneous teams becomes stronger in times of ethnic tensions in the country. Similarly, within the context of the garment industry, Kato and Shu [2011] use data from a Chinese garment factory to show that the effect of team incentives to increase productivity depends on the composition of work teams out of urban and rural migrant workers. Furthermore, Hamilton et al. [2003] study the introduction of team work in a U.S. garment factory, which has a similar set-up as the factories I study in Bangladesh, and find a significant increase in productivity from team work.[2]

This paper proceeds as follows. The next section introduces more background information about the factories and describes in more detail the dataset collected, while section three presents the non-experimental results on productivity spill-overs,

[2]Further studies in this field include Shi [2010] and Lazear [2000], who find strong and very similar increases in labor productivity at two companies when they switch their wage scheme from hourly to piece-rate pay. Mas and Moretti [2009] study peer effects among cashiers in a supermarket in the US, and whether altruism or social pressure makes people work harder when their work effort has positive externalities on the work pressure of co-workers. They mainly find the latter effect to be at work. Das et al. [2013] investigate the effect of worker training on shift productivity in an Indian steel mill. Nagin et al. [2002] could randomly vary the supervisory monitoring rate at a telephone solicitation company, to investigate how this affected the shirking behaviour of employees.

using the whole collected dataset. Section four provides more details on the experiment and shows its main effects. Section five presents results on the interplay of the intervention with social ties, while section six will conclude.

3.2 Background and Data

This study was conducted at three large garment factories in Bangladesh, which has emerged as the third largest garment exporter in the world over the last years.[3] For local industry standards, the three factories are very large and modern. Both ownership and management are domestic, and all output is produced for the export market. The factories produce mainly t-shirts, polo shirts, shirts and pants. The factories vary in size, with between 1,200 - 5,000 workers employed in their sewing departments.

Table 3.1 provides key characteristics of the three factories in the sample. Factory 2 is smaller than the other two factories, with only 17 sewing lines located on four sewing floors. However, it has a higher number of workers per line. Factory 3 is on the other side of the spectrum in many respects. It has many more lines than Factory 2, which, however, have on average less than a third of the number of workers per line. However, in most cases, two lines share one line chief, and in some cases even four lines share one. Factory 1 lies in-between the other two factories on most dimensions. It has 59 lines on six different sewing floors, each line with its own line chief, and the size of the lines being closer to the ones from Factory 3. Workers on lines are faced with new garments on average every 16 days in Factory 1, while roughly every 8-10 days at Factory 2 and 3.

Sewing lines are organized as assembly lines in which each worker only does one sewing operation, then passes on the garment to the next operator who does another sewing operation. Additionally each line has one to three quality inspectors, and garments found with quality defects that cannot easily be rectified are sorted out and not counted in the line-wise output measure. The main tasks of the line chief is to respond to problems from operators, to instruct operators with new tasks when a new garment is started, and generally to keep work discipline and productivity high. Both workers (including line chiefs) and orders are allocated to sewing lines from the factories' central planning departments. Workers have fixed places at lines, which they usually only change if promoted to new positions. Only when

[3]Source: WTO, International Trade Statistics 2014: www.wto.org/english/res_e/statis_e/its_e.htm

Table 3.1: Factory Characteristics

	Factory 1	Factory 2	Factory 3
Nbr. Sewing Floors	6	4	14
Nbr. Sewing Lines	59	17	183
Nbr. Workers in Sewing Section	ca. 2000	ca. 1200	ca. 5000
Nbr. Workers in whole Factory	ca. 5000	ca. 2000	ca. 9000
Nbr. Buyers	28	74	10
Nbr. Garments in Data	866	839	1048
Avg. Nbr. Lines /Garment	3.12	1.49	3.94
Avg. Nbr. Days /Garment & Line	16.4	9.5	8.5
Avg. Nbr. Workers /Line	30.9	72.2	23.2
S.Dev. Nbr. Workers /Line	8.0	10.8	5.2

Notes: All information from production data collected from factories, except for 'Nbr. Workers in ...' which is from surveys of factory management.

absent workers on other lines need to be replaced do workers occasionally switch to different lines on a day-by-day basis. However, workers with production experience on some garments are generally not reallocated to other lines if these lines also start producing the same garment. Thus, it is unlikely that such reallocations of workers drive the observed productivity spill-overs across lines producing the same garment. Appendix C furthermore presents the results of a placebo test which also indicates that worker movements are not behind the productivity increases of later lines producing the same garments.

The main dataset used for the analysis contains line-wise production data for all lines in the factories for 30 consecutive months from Factory 1 and 2, and for 8 consecutive months from Factory 3, as this factory was recruited for this project only at a later point in time. This dataset includes daily sewing line efficiency as calculated by the factories, the garment being produced by a line on a given day, its buyer, and the Standard Minute Value ('SMV') of the order. The SMV is a garment-specific value, calculated prior to the start of production of a garment. It is the sum of the times in seconds it takes to perform each sewing operation to assemble one piece of the garment, providing a measure of the required labour input per piece under ideal production conditions. To calculate the line efficiency measure, daily piecewise output is multiplied by the garment-specific SMV, and then divided by total labour input on that line and day measured in worker-minutes. As the SMV is also essential in price negotiations with the buyers, the calculation of the SMV

Table 3.2: Line Chief Characteristics

	Factory 1	Factory 2	Factory 3
Avg. Age	30.0	30.0	32.2
S.Dev. Age	3.8	5.8	5.6
Avg. Years working in Factory	4.3	5.0	6.2
Avg. Y.s working as LC in Fact.	2.4	1.8	3.5
Avg. Y.s working as LC on current line	1.2	1.5	2.1
Promoted internally	58%	73%	52%
N	53	15	60
N Female LCs	1	1	0

Notes: All information from survey of all line chiefs in factory. Promoted internally is percentage of current line chiefs who worked on lower position in same factory before and were subsequently promoted to line chief.

is done professionally, and its breakdown into the individual sewing operations is being scrutinized by the buyer. Therefore, the efficiency measure is of high quality and comparability across different garments.

I also conducted a survey of all line chiefs working at the three factories, which collected demographic and career information, plus information on the social networks between all line chiefs on six dimensions: kinship, knowing each other since before working at the factory, having worked together at another factory before, having visited each other's home, spending lunch breaks together, and generally 'being friends'. This network data is discussed in more detail in chapter four of this thesis. Table 3.2 presents summary statistics from the surveys. Line chiefs are around the age of 30 at all factories, and only 2 out of 128 line chiefs interviewed were female.[4] They have worked as line chiefs on average for 1.8-3.5 years at the factories, and report to be line chief of the line they were at the time of the survey for already more than one year on average. This also fits with the accounts from the factory management and the production data; line chiefs generally have a fixed line and are only rarely reallocated. At all three factories, line chiefs report to have on average ca. 10 years of schooling, which is equivalent with the Bangladeshi Secondary School Certificate (SSC).

[4]The average share of female workers on the sewing lines is ca. 80%. Workers typically start working in the garment industry at the age of 18 (child labour regulation these days being enforced through foreign buyers), and stop by the age of 25-30, unless promoted to quality control, mechanic or supervisory positions. However, only very few women get promoted to these positions.

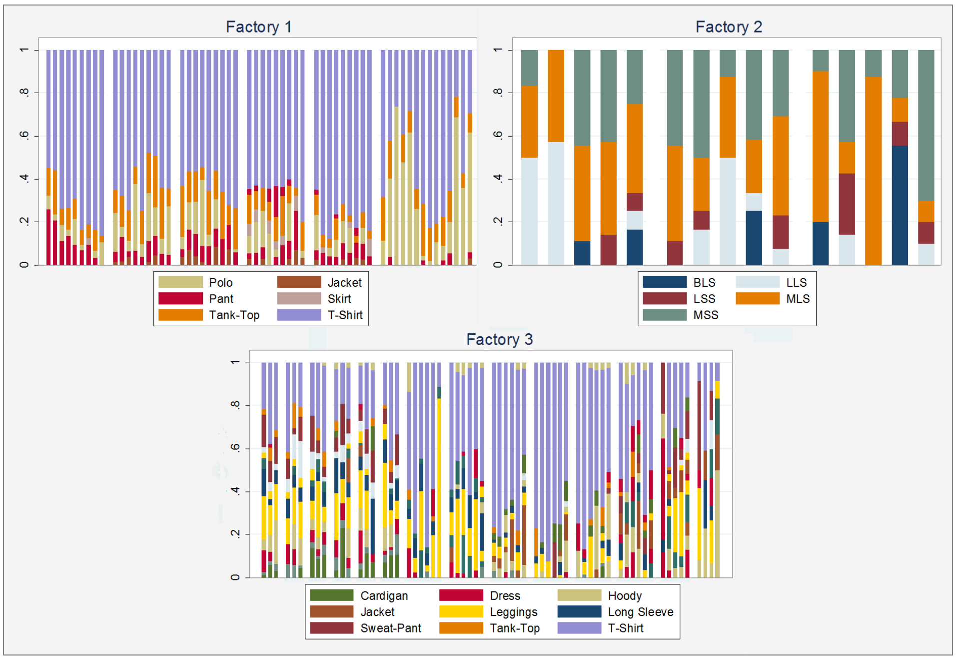
Sewing lines are kept homogeneous in terms of size and productivity within the factories by the management, and workers are not sorted to lines according to experience or productivity.[5] The reason for this lies in the high flexibility required in operations. Buyers place orders with low predictability and close delivery deadlines, and frequent disruptions to the production process (power failures, unrest outside factories, problems in supply and delivery chains, missing inputs) often require re-allocations and re-prioritization of orders to lines. Therefore, it is not optimal to have differentiated lines specializing on certain types of garments. This non-specialization of lines on certain types of garments can also be seen in the stacked bar charts for each of the three factories in Figure 3.2, in which each bar represents a sewing line, and the wider spaces between the bars separate sewing lines located on different floors. The differently coloured parts of the bars represent the shares of different garment types (e.g. t-shirts, polos, pants,) among all garments the lines produce. While some variation can be expected, in general the graphs show little patterns of lines specializing on certain types of garments.

Lines also do not specialize in whether they are typically the first or a later line in the roll-out of garment orders across lines. Figure 3.3 shows similar stacked bar chart as Figure 3.2, but this time the differently coloured parts of the bars indicate the share of the garments the line produced for which it was the first (orange), the second (light blue), or the third or later line (dark blue) to produce it in the factory. Again, few obvious patterns of lines being more often allocated garments early on or later can be seen.[6] According to the production engineers in the planning departments, incoming orders are prioritized based on the importance of the buyer to the firm and how close the delivery date is, and are then essentially allocated to the 'next free line'. This speaks against the higher productivity of lines producing garments that have already been produced on other lines before being driven by selection of certain types of lines with higher productivity into usually producing a garment not as first line in the factory, but at a later time.

[5] An exception are the 'sample lines' on which first samples are produced for buyers during the negotiations process for new orders, on which often the most experienced workers work. However, sample lines are not included in my dataset.

[6] At Factory 3, the six floors to the left of the graph produce for one large buyer, while the other floors for other buyers. As this buyer places larger orders, which are produced on average on more lines, the lines on these floors are on average less often the first to produce a given order, and more often the second, or later line.

Figure 3.2: Garment Types produced on different Sewing Lines



Graphs represent the types of garments produced by different sewing lines at the three sample factories. Each bar in the graphs represents a sewing line, and the wider spaces between bars separate sewing lines from different sewing floors. The differently coloured stacked parts of the bars represent different types of garments that the lines produced. Legends show colours for most common garment types only, for illustration. In sub-graph of Factory 3, each bar represents a line chief instead of a line (some line chiefs at this factory look after 2 or 4 lines), to keep the number of bars in graph parsimonious. Graph shows types for only 15 out of 17 lines for Factory 2, as type data is missing for two lines.

3.3 General Evidence on Productivity Spill-over

Before turning to the results from the randomized intervention, this section explores in the overall production dataset to what extent line productivity profits from output of the same garment already produced on other lines before. Given the evidence from Figure 3.3, lines do not specialize in whether they typically are the first or a later line in the factory to produce a garment. Therefore these effects should not be driven by selection of higher productivity lines into producing garments later in the roll-out process of garments across lines. The identifying assumption is then that also for types of garments, which lines are for some reason better or worse at producing, they do not specialize into producing those garments early on or at a later stage in the roll-out process.

I observe 1,257 garments in the overall dataset which have been produced

Figure 3.3: Start Rank of Garments produced on different Sewing Lines

Graphs show for each sewing line in the three factories for which share of the garments they produce the lines are the first (orange), the second (light blue), or the third or later line (dark blue) to produce that garment in the factory. Each bar represents one line, and the wider spaces between bars separate lines located on different sewing floors. In the sub-graph of Factory 3, each bar represents a line chief instead of a line (some line chiefs at this factory look after 2 - 4 lines), to keep the number of bars in graph parsimonious.

on more than one sewing line in the factories. In total, there are 4,964 instances of a sewing line staring to produce one of these garments (from now on I will refer to the event of a line starting a new garment as a 'garment start'). Therefore, these garments are produced on average on 3.95 different lines. Figure 3.1 showed that on average sewing lines reach their long run productivity level again five days after starting a new garment. Therefore, I keep the daily line productivity observations from the first five days a line produces a new garment for the sample of the regressions. Denote the n'th day a sewing line produces a garment as the n'th 'garment-day'. Thus the sample consists of all observations with garment-day less or equal five.

The basic econometric model I estimate in this section is of the following form:

$$y_{fignt} = \sum_n \beta_n^A ln(A_{ign}) + \sum_n \beta_n^S ln(S_{ign}) + \sum_n \beta_n^X X_{gn} + \alpha_{fin} + \gamma_{fn} + \epsilon_{fignt} \quad (3.1)$$

Productivity y_{fignt} of sewing line i in factory f producing garment g in week t on its n'th garment-day is regressed on the output A_{ign} of the same garment that has already been produced on all other sewing lines in the factory up to, but excluding, the day on which line i started producing garment g. I interact this previous output from other lines with fixed effects for garment-day n. Thus, the effect of previous output of the same garment is estimated separately for each of the five garment-days included in the sample, to see for how long previous output affects productivity of a new line producing the same garment. I use the log of previous output of the same garment as I expect each additional produced piece of the garment to have a diminishing marginal effect on the stock of knowledge with the garment held by other workers.

I additionally include in the regression the output S_{ign} of the same garment produced on all other lines on the same sewing floor, with its effect estimated again separately for each garment-day. Sewing lines in the three factories are bundled on sewing floors which contain on average 5-10 lines, and sewing lines located on the same floor are running parallel and only 2-3 meters apart from each other. Sewing floors, on the other hand, are either located on top of each other in the same building, or in different buildings. Therefore, to get from a sewing line on one floor to one on another requires at least leaving one's line out of sight and calling distance. Furthermore, each sewing floor typically has its own floor production manager, who could transfer knowledge with a garment he gained on one line to other lines on his floor.[7] For these reasons we could expect a priori the effect of production experience with the same garment gained by lines on the same floor to differ from the effect of experience gained by lines on other floors.

I control for fixed effects α_{fin} on the line chief - garment-day level.[8] Thus, I estimate the effect of previous output of the same garment (on the same floor) as deviation of line productivity from learning curves estimated separately for each line chief. X_{gn} is a vector of garment characteristics interacted with garment-day, which in the baseline estimation includes the SMV of the garment, and individual dummy variables for each type of garment. Finally, γ_{ft} are time fixed effects on the factory-week level. Standard errors are clustered on the line chief level.

[7] I use 'he' and 'his floor' as all floor managers in the three factories are male.

[8] The later results on the experimental intervention use line chief fixed effects, as the intervention treats line chiefs. Thus, for consistency, I also use line chief fixed effects in this section. All results are qualitatively similar when using line fixed effects instead of line chief fixed effects.

Column 1 of Table 3.3 shows the results from estimating the empirical model from equation 3.1. The 'All Other Lines, Day n' coefficients represent the effect of log previous output of the same garment on all other lines in the factory for each of the first five garment-days, while the 'Lines Same Floor, Day n' coefficients represent the effect of log prior output from other lines on the same floor. Output on any other line increases productivity of later lines, and its effect does not seem to reduce with the number of days the line already produces the garment. If including more garment-days into the sample, the effect diminishes only slowly, but becoming increasingly insignificant. On the other hand, output from lines on the same floor has a large additional effect above and beyond the effect of previous output on any line, which however disappears after the third production day.

The non-diminishing effect of output on any other line on productivity disappears, however, when including additional fixed effects on the level of the 1,257 individual garments, as shown in column 2 of Table 3.3. This could point to selection of garments with potential for higher productivity, which is not captured by garment type and SMV, into being produced on a larger number of lines. However, this hypothesis is not supported by column 3, which repeats column 2, but excludes those instances of lines starting garments in which no other line has produced the garment before. The effect of previous output is now estimated only off its intensive margin, off the amount of the garment that was already produced on other lines before, and not anymore off the extensive margin, whether another line has produced the garment at all or not. Now both previous output on all other lines shows a significant effect, as does the interaction with whether the garment has been produced on the same floor before. The picture which now emerges is that the main effect of previous output does not fundamentally depend on whether the garment has been produced on the same or another floor before, with the effect being about a third larger if the garment has been produced on the same floor before. The relatively low level of this additional effect speaks against the hypothesis that the effect of previous output is mainly a 'within-person effect', in the sense that the floor level production manager gains knowledge about a certain garment on one line on his floor and then applies the knowledge on another line. It seems more likely that knowledge on a certain garment is gained by workers on one line, and then communicated to workers on another, either on the level of ordinary workers, line chiefs, or floor supervisors.[9]

[9]In principle, the effect could still be a within-person effect in that the production head or head engineer of the whole factory learns about all garments produced on any line in the factory and then applies his or her knowledge when other lines anywhere in the factory start the same garment. However, between two lines (at Factory 2) and 15 lines (Factory 3) on average change the garment

Table 3.3: Non-Experimental Results

	(1)		(2)		(3) Start-Rank > 1		(4) Start-Rank 1		(5) Start-Rank > 1	
VARIABLES	Efficiency	SE	Efficiency	SE	Efficiency	SE	Efficiency	SE	Efficiency	SE
PRIOR OUTPUT										
All other Lines, Day 1	0.268**	(0.12)	0.092	(0.12)	0.898***	(0.23)			0.892***	(0.23)
All other Lines, Day 2	0.275***	(0.09)	0.108	(0.08)	0.636***	(0.24)			0.635***	(0.24)
All other Lines, Day 3	0.215***	(0.08)	0.951	(0.09)	0.137	(0.22)			0.177	(0.22)
All other Lines, Day 4	0.328***	(0.08)	0.215**	(0.09)	0.210	(0.21)			0.286	(0.21)
All other Lines, Day 5	0.179**	(0.07)	0.131	(0.08)	0.235	(0.24)			0.319	(0.24)
Lines Same Floor, Day 1	0.567***	(0.14)	0.426***	(0.12)	0.301**	(0.12)			0.421***	(0.12)
Lines Same Floor, Day 2	0.284**	(0.11)	0.213**	(0.09)	0.208*	(0.11)			0.289***	(0.10)
Lines Same Floor, Day 3	0.235**	(0.10)	0.163*	(0.08)	0.248**	(0.10)			0.289***	(0.10)
Lines Same Floor, Day 4	0.007	(0.09)	-0.058	(0.07)	-0.052	(0.10)			0.004	(0.10)
Lines Same Floor, Day 5	0.097	(0.09)	-0.005	(0.07)	-0.005	(0.11)			0.039	(0.11)
OUTPUT FIRST DAY										
All other Lines, Day 1							2.475***	(0.60)	0.980***	(0.31)
All other Lines, Day 2							1.416***	(0.42)	0.473**	(0.23)
All other Lines, Day 3							0.958**	(0.37)	0.069	(0.22)
All other Lines, Day 4							0.667*	(0.40)	0.508**	(0.25)
All other Lines, Day 5							0.352	(0.29)	0.313	(0.27)
Lines Same Floor, Day 1							1.462**	(0.62)	0.634*	(0.34)
Lines Same Floor, Day 2							1.158**	(0.46)	0.717***	(0.25)
Lines Same Floor, Day 3							0.505	(0.39)	0.690***	(0.21)
Lines Same Floor, Day 4							0.146	(0.43)	-0.011	(0.25)
Lines Same Floor, Day 5							0.122	(0.34)	0.110	(0.29)
Constant	39.459***	(3.86)	2,357.4***	(559.2)	18.317**	(7.11)	25.950***	(5.16)	29.585***	(4.72)
Observations	16,169		16,169		9,697		6,449		9680	
R^2	0.481		0.481		0.701		0.614		0.709	
LC-Grmt.Day FE	YES		YES		YES		YES		YES	
Grmt.Type-Grmt.Day FE	YES		YES		YES		YES		YES	
Garment FE			YES		YES				YES	
SMV * Grmt.Day FE	YES		YES		YES		YES		YES	
Week-Factory FE	YES		YES		YES		YES		YES	

Columns 1-3 show the results from regressing daily line productivity from first five days a line produces a new garment on the log of output of the same garment produced on all other lines in factory up to, but excluding, the day the line starts producing it as well. This previous output is interacted with fixed effects for garment-day ('All other Lines, Day n'). Similarly, log previous output of the garment on all other lines on the same sewing floor, interacted with garment day, is included separately ('Lines Same Floor, Day n'). Regressions control for line chief - garment-day fixed effects, garment type - garment-day fixed effects, SMV interacted with garment-day, and week - factory fixed effects. Column 2,3 and 5 also include fixed effects for the individual garments. Column 3 only includes observations in which a positive amount of the garment has been produced on other lines before. Column 4 regresses instead daily line productivity on the first five garment days from the first lines which produce a certain garment in the factory on log output of the garment from lines which also started producing the garment on the same day, interacted with garment-day ('All other Lines, Day n'), and separately from lines located on the same floor ('Lines same Floor, Day n'). Column 5 regresses output on both previous output and on output from other lines starting to produce the garment on the same day. Standard errors clustered on the line chief level: *** p<0.01, ** p<0.05, * p<0.1.

So far, I showed the effect of previous output of the same garment on productivity of new lines starting to produce the same garment. Most of the literature on organizational learning used such effects as evidence for learning and knowledge exchange in the factories. However, such effects could also be driven by other peer effects. The mere fact that other workers in the factory produce the same product could increase productivity even without learning effects. Workers could compete about who is most productive with a product, or the productivity of some workers could provide the factory management with a benchmark against which it could compare productivity of other workers, and therefore more easily find out if workers slack. To find more evidence on whether indeed knowledge transfers drive the productivity increases of later lines producing the same garment, I run a simple placebo test. If the increases in productivity that we observed would be driven by competition or less slack, these effects should arguably be even more prevalent if lines start producing the same garment on the same day, as in these cases, the playing field for competition or comparing efficiency should be more levelled. Furthermore, one can study the effect of lines starting producing the same garment on the same day in those cases when sewing lines are the first lines in the factory to produce a certain garment. In these cases, no other workers have already gained experience with the garment which could be shared. Therefore, learning effects should be absent. Thus, in column 4 of Table 3.3, I study the productivity of the first lines in the factory to produce a certain garment, and check whether it is increased if more than one line start producing it on the same day, and if this effect is stronger if the other lines starting the garment on the same day are located on the same floor.[10] And, indeed, strong and significant such effects can be seen in column 4.

This effect among first lines could be driven by selection of certain garments into instances where more than one line is the first to produce it, such as rushed starts, when an order needs to be completed quickly, and more than one line therefore start producing it on the same day. Furthermore, as each garment is only started once for the first time in the factory, I cannot use garment fixed effects. Thus, the

on any given day. Thus it seems unlikely that these two persons could carry within them the whole knowledge exchange process (while the additional effect of the floor supervisors, who are involved in much less garment changes, is much smaller).

[10]I regress efficiency on log output produced on all other lines on the first day of production which also started to produce the garment on the same day. There are 2,002 instances in the data of a line starting a garment that has not been produced before on another line. At 1,170 of them, some other line in the factory also starts producing the same garment on the same day, and at 1,054 instances, at least one of these other lines is located on the same floor.

effects could be confounded by selection of garments with certain characteristics to multiple first lines which is not controlled for by their type or SMV. Finally, the presence of the effect among the first lines does not automatically imply that the effect of previous output among later lines is not driven by learning effects; the two effects could be explained by different mechanisms.

To gain additional insight on whether the effect among the first and later lines are driven by the same or different mechanisms, I conduct a horse-race among the two specifications in column 5, which replicates the specification from column 3, but adds log output produced by lines starting the garment on the same day as additional independent variables. Note that this output is not mechanically co-linear with previous output of the same garment, as previous output includes output only up to, but excluding, the day the line also starts producing the same garment. Thus, it cannot include output from lines starting the garment on the same day. In fact, previous output and output from lines starting the garment on the same day is negatively correlated, both overall and within units. As in this specification I can observe again multiple starts per garment on different lines, I include garment fixed effects again. The results from column 5 show that in this specification, both effects are present. Lines are more productive the more of the same garment has already been produced previously on other lines, and the more is being produced on the same day on other lines.

To conclude this section, the existence of spill-over effects among first lines should caution against interpreting the effects of output on other lines as learning, even though in a regression which attempts to incorporate both peer and learning effects, both effects seem to persevere. However, even when controlling for output from other lines on the same day, previous output could still capture other possible forms of peer effects. Therefore, the next section presents the results of a randomized intervention which introduced exogenous variation in the likelihood that production knowledge on garments is communicated, and shows that this intervention did have an effect on line productivity when starting new garments.

3.4 Randomized Help Provision

To identify the effect that knowledge exchange between co-workers has on productivity growth, I carried out a randomized management intervention at the three sample factories. Whenever a line on randomly selected 'treatment' sewing floors began producing a garment that had already been produced by *any other line chief in the factory*, the most senior line chief with previous experience on the garment was instructed by the factory's production management to brief for 15-30 minutes the line chief without the garment-specific experience on how initial production problems with the garment were overcome on the earlier line.

The intention of this treatment was to exogenously increase the potential for knowledge exchange on the production process of the garment between randomly selected pairs of line chiefs, by lowering the costs of helping. In particular, the intervention can be thought to decrease two parts of the cost of seeking and providing help. First, the possible perceived cost to approach someone else for help, as one exposes a lack of knowledge on how to solve certain production problems (Lee [2002]; DePaulo and Fisher [1980]). By having someone else being sent by higher ups to share his or her experience with the garment, knowledge is shared without an initial request for help which would reveal a lack of knowledge. Second, especially if the person asked to share his or her knowledge needs to go to the workplace of the other person to provide effective help, help provision can be thought to have a fixed and a variable cost component. Allowing the distraction of listening to someone's request for help, and possibly moving to the other person's workplace would constitute a fixed cost. Once this cost is borne, one would need to decide how much effort to spend into analysing the problem at hand and figuring out an effective way to communicate a possible solution, which introduces a variable cost component into help provision. While the randomized help provision does not eliminate either cost, it does make the fixed cost of engaging with the other worker and moving to his or her workplace sunk, as the worker cannot decide anymore whether or not to bear these costs. Thus we can think of this cost as being taken out of the cost-benefit analysis of the worker asked to provide help, when deciding whether to do so.

The experiment ran on the treatment floors for four months, from June to September 2014. The production data shows 377 non-first garment starts on the treatment floors during this time at which some other line chief had already produced the same garment, which should have been treated in the randomized experiment. The treatment protocol was implemented by the industrial engineers from the factories. The engineers were provided experimental log books to record each instance of a treatment of a garment start. According to these logbooks, 220 treatments were actually administered during this time, of which 125 could be matched with garment starts in the production data.[11] However, it is likely that compliance was higher than indicated by these numbers. The implementing engineers admitted underreporting of treatments in the logbooks. Among the actually treated garment starts, there is likely to be selection into treatment of garment starts for which the treatment was expected to have a stronger effect.[12] For these reasons, the analysis will focus on the intention to treat effect, assuming that any start of a garment that should have been treated was actually treated. Any non-compliance with the treatment would lead to an underestimation of the treatment effect.[13]

The sample from which floors were randomly selected consisted of 17 floors across the three factories (Factory 3 requested to include only six out of its 14 floors in the sample).[14] Randomization across floors was chosen to make compliance with the randomized implementation as simple as possible for the factory management. The original intention was to randomize treatment across sewing lines. However, the factory managements were worried that it would be too difficult for their staff to remember which lines should be treated and which not. Furthermore, there was the concern that if the intervention is implemented at some lines on a given floor, and if it proves effective, its implementation would quickly be copied by other lines in the same floor, which usually operate just a few meters away. Table 3.4 below shows tests of balanced outcomes over observable average line and line chief characteristics from April and May 2014, just before the start of the intervention, when the random selection of units was done. No observable line chief or line characteristics differ significantly across treatment and control floors, except for average line

[11] At 16 out of these matched reports, however, according to the production data, no other line chief had produced the garment before. It seems that in these instances, line chiefs that had already produced similar garments were sent to give instructions. The main reason for non-matching was mismatch of the garment identifier provided in the logbooks, which could not be matched to any garment the line was producing according to the production data in the days around which the treatment was reportedly done.

[12] All three factories reported that already before the intervention, at times they sent line chiefs to other lines to help their co-workers with new garments, if they are already experienced with the garments. However, this behaviour was not institutionalized in any of the three factories. To the extent that the factories already induced line chiefs to help each other, the factories were instructed to not change their behaviour on the control floors, while always sending line chiefs to brief others starting the same garments on the treatment floors.

[13] Note that there is no indication of garment starts on control floors being treated. The logbooks do not show any such treatment, and also the production managers who implemented the intervention showed no sign of confusion about which garment starts should be treated, and which not.

[14] In fact, the sample of floors over which the randomization occurred consisted of only 15 sewing floors. However, at factory 1 and 2, one floor was randomly chosen at each factory, and one (physical) half of the floor randomly selected into treatment. Therefore, the randomization occurred effectively across 17 units, 13 full floors, and 4 half floors.

Table 3.4: Balancing of Randomization across Sewing Lines

	Control Lines	Treated Lines	N
Line Chief Characteristics			
Age	30.06	0.16	73
Seniority Factory (months)	61.79	2.48	74
Seniority as Line Chief (months)	34.71	0.12	74
Seniority as LC current line (months)	9.45	6.54	71
Promoted Internally	64.7%	-0.02%	74
Nbr. Social Connections	2.66	0.47	74
Education	15.27	-0.22	72
Line Characteristics			
Avg. Number Worker	28.92	1.59	140
Avg. Daily Hours	9.57	0.16	140
Avg. Efficiency	53.53	-3.09*	140
Avg. SMV	10.76	-0.81	140

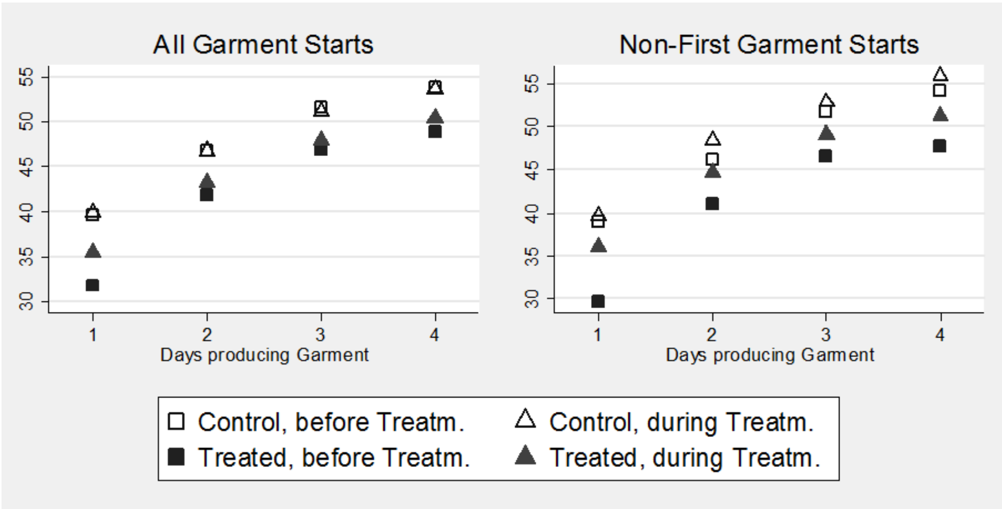
Notes: Line Chief characteristics from line chief surveys. Line characteristics from production data. Values in 'Treated Lines' columns show deviation of average values from treated lines from those from control lines, with * reflecting statistical difference on 10% level.

efficiency, which was lower on the treated floors (p-value 0.082).[15]

As the intervention was conducted at the end of the time covered by the collected production data, a substantial amount of pre-intervention data is available. Figure 3.4 plots the average efficiency over the first four days a line produces a new garment for four different cases: treated lines before and during the intervention, and non-treated lines before and during the intervention in the factory. I use data from the beginning of 2014 until end of September 2014, when the initially defined treatment time ended. The left hand side shows the average learning curves over all garment starts, while the right hand side shows the graph when only considering non-first garment starts, which are in principle 'treatable'. Prior to the start of the intervention, and compared to control lines, treated lines had on average lower efficiency values at the first days a new garment was produced, which fits with the overall lower efficiency among lines in treatment floors shown in Table 3.4. This difference is not accounted for by observable characteristics of lines or line chiefs,

[15] All differences, except for line efficiency, remain insignificant when controlling for factory fixed effects.

Figure 3.4: Pre-and Post- Treatment Learning Curves

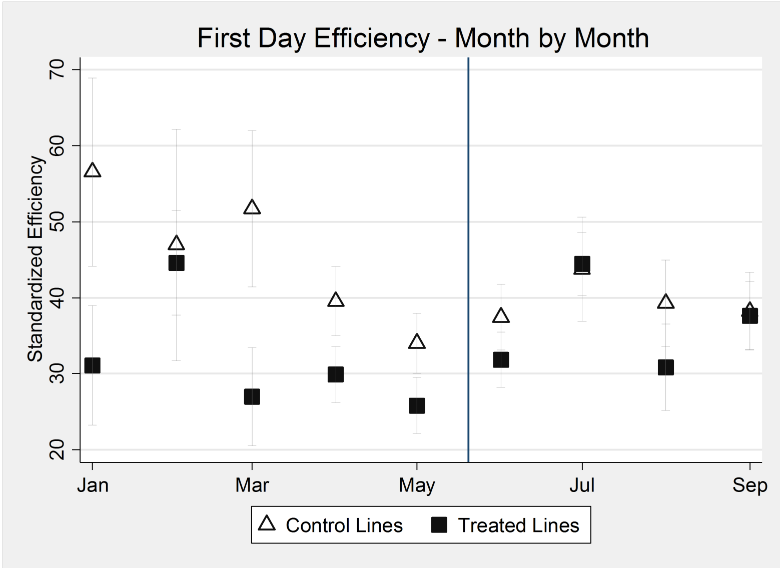


Both graphs plot average efficiency over the first four days a line produces a new garment for four different cases: From treatment floors, prior to start of treatment (solid square symbols), from treatment floors, during experiment (solid triangle symbols), from control floors prior to start of treatment (hollow square symbols), and from control floors during time of treatment (hollow triangle symbols). Left hand graph uses sample of all garment starts, while right hand side only from non-first garment starts. Efficiency standardized on factory-level.

and is driven by two out of the three factories. However, as shown in Figure 3.4, while productivity remains constant across pre-treatment and treatment time on control floors, treated lines experience an upward shift in their learning curves during the time of the treatment. Interpreting the results in a difference-in-difference framework, the intervention indeed had an effect in raising efficiency, especially at the first day a new garment was produced.

Using a difference in difference framework to identify treatment effects, one should check whether pre-trends differ between treated and non-treated units. Figure 3.5 provides such a check. For the first day a line produces a new garment that has already been produced by another line before, it plots monthly average efficiency over the year 2014, separately for lines selected for treatment (square symbols) and as control lines (triangle symbols). The vertical line indicates the start of the treatment with June 2014. It indeed looks as if first day efficiency was systematically lower at floors selected for treatment in the months before the start of the treatment. This difference is then greatly reduced with the onset of the intervention due to an upward shift of first day efficiency on treatment floors, especially when compared to the three months directly preceding the start of the intervention.

Figure 3.5: Pre-Post Intervention Start Trends for First Day Efficiency



Graph shows average monthly efficiency of lines on the first day they start producing a new garment, separately for lines selected for treatment (solid squares) and lines not selected (hollow triangles). The vertical line indicates start of treatment from June 2014 on. Capped bars represent 95% confidence intervals.

To estimate the intention-to-treat effect of the intervention in a difference-in-difference approach, I keep, similar as above, the observations from the first three garment-days from each garment start from those floors which were part of the sample among which treatment floors were selected, from January until September 2014. As the intervention ran in the treatment factories from June to the end of September 2014, it covers roughly the second half of the sample.[16] Using this sample, I run the following baseline regression:

$$y_{fignt} = \sum_n \beta_n^T Treat_{ign} + \alpha_{fin} + \gamma_{fn} + \delta_{ign} + \epsilon_{fignt} \tag{3.2}$$

Productvity y_{fignt} of line i producing garment g on one of the first three garment-days n is regressed on a dummy $Treat_{ign}$ for the start of the garment randomly being selected for treatment, interacted with fixed effects for the three

[16] More precisely, the intervention started on 21^{st} May 2014 at Factory 2, on 23^{rd} May in Factory 1, and on 1^{st} June in Factory 3

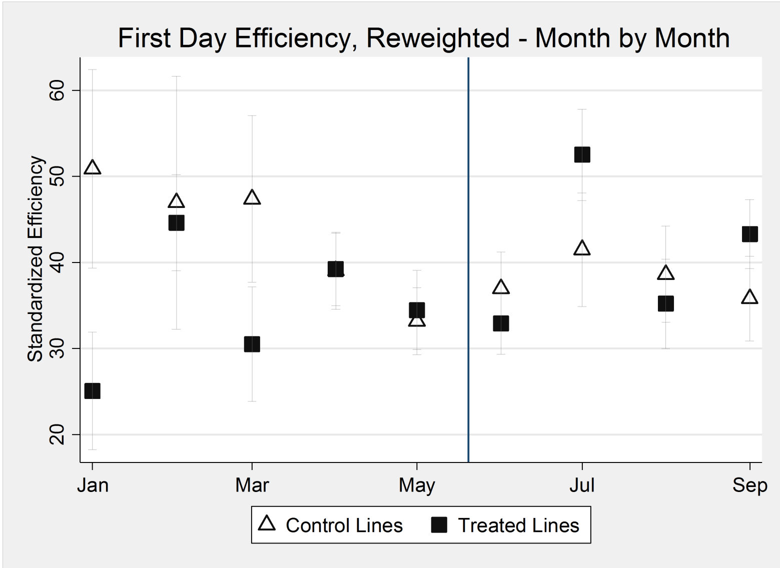
different garment-days included in the sample. I control, as in the previous section, for fixed effects α_{fin} on the line-chief - garment-day level, and γ_{ft} on the factory-week level. Furthermore, I include fixed effects δ_{ign} for the 'rank' of the line in the roll-out of the garment, which indicates how many other line chiefs in the factory have already produced the garment, interacted with garment-day fixed effects.[17]

The reduction in the difference of starting productivity with the onset of the intervention could imply that the results are caused by some other form of catch up of productivity on treatment floors relative to control floors, which coincided with the start of the intervention. To address this concern, I apply the reweighting approach by DiNardo et al. [1996] to all regressions. It reweights observations from the treatment floors such that in the pre-treatment time the average learning curves do not differ anymore between treatment and control floors. I use the approach in a similar way as Duflo et al. [2013], who adapted it to control for possible endogenous selection into treatment. Their basic idea is to reweight observations from a controlled experiment such that independent variables that were not balanced pre-treatment between treated and control units become balanced after the reweighting. In this paper, I apply this approach to correct for the fact that the dependent variable of efficiency on the first days a line starts a new garment is not balanced between treatment and control groups prior to the start of the randomized experiment. Identification using the reweighting approach relies on the assumption that the treatment effect does not depend on the distribution of the independent and dependent variables, as the approach creates artificial counterfactual distributions in the sample used to estimate the treatment effect. More details on the implementation of the approach are shown in Appendix D.

Figure 3.6 demonstrates the reweighting approach, replicating Figure 3.5 after reweighting the data such that efficiency on the first day a line produces a new garment which has already been produced on another line before is balanced in the two months before the start of the intervention, the same time frame used to create the results for the balancing tests from Table 3.4. Indeed, the graph now shows that in the reweighted sample, efficiency on the first day a line starts producing a new garment that has already been produced on another line before does not significantly differ anymore between treatment and control floors in the two months before the

[17] Instead of rank fixed effects, I could have also controlled for cumulative output of the same garment on previous lines, the central variable of interest in the regression from the previous section. However, to more flexibly control for how many lines have already produced the garment before, I instead use rank fixed effects (interacted with garment-day).

Figure 3.6: Pre-Post Intervention Start Trends for First Day Efficiency, Reweighted Data



Graph shows average monthly efficiency of lines on the first day they start producing a new garment, separately for lines selected for treatment (solid squares) and lines selected as controls (hollow triangels), with efficiency data reweighted following the approach from DiNardo et al. [1996]. The vertical line indicates start of treatment from June 2014 on. The lines through the symbols represent 95% confidence intervals.

start of the intervention. In the reweighted data, no effect of the treatment is visible during the first month of the intervention, in June 2014. A positive effect is now mainly visible during July and September 2014.

Due to the small number of only 17 clusters over which the randomization was conducted, special attention needs to be given for inference, as even standard errors clustered on the 17 floors can be biased downwards, as shown by Cameron et al. (2008). They suggest wild cluster bootstrap to estimate adequate standard errors, which will be applied at all regressions estimating the effects of the randomized intervention.

Column 1 from Table 3.5 shows the results when estimating the model from equation 3.2, using the reweighted data. A significant positive effect on productivity on the first day of production of a new garment can be seen. First day efficiency is increased by 4.09 efficiency units, which resembles 19.2% of the standard deviation of first day productivity. Average first day productivity of lines if other lines have already produced the garment is 40.0 efficiency units, and overall long run efficiency is 50.1. Thus, the intervention reduces the average gap of first day to long run efficiency by about 40%. The effect becomes successively smaller and insignificant on the second and third day of production. Column 2 adds garment fixed effects, which had been shown in Table 3.3 to be important when estimating the general effect of previous cumulative output on productivity. In principle, the characteristics of the garments produced on treatment and control lines should be balanced, due to their random selection, and Table 3.4 shows no significant difference in the SMV of garments across these two types of lines. However, the non-balanced pre-treatment efficiency levels between treatment and control lines could be due to different garments being produced on these floors whose effect on efficiency is not captured by their SMV, and the treatment effects we see could be induced spuriously by a change in the garments produced on treatment lines. Thus, column 2 includes garment fixed effects, which makes the estimate of the effect on first day efficiency somewhat larger and more significant.

So far, I included all instances of a line starting a new garment in the sample. However, only those starts of garments at which at least one other line has already produced the same garment before can be treated, as the design of the intervention requires the presence of one line chief who is already experienced with the garment, who can administer the briefing. Thus, a more direct way of estimating the intention-to-treat effect is to restrict the sample only to the non-first garment starts, both before and after the start of the intervention at the factory, and on treatment and control floors, which is done in column 3. The results remain qualitatively the same when estimated only on this restricted sample.

While garment fixed effects control for all unobservable characteristics of the garment, they do not capture possible interaction effects of garments and lines. And while lines are in principle not specialized on certain garment types, they could nevertheless be differentially productive for different types of garments, for example as they happened to have produced more of a certain garment type in the past than another. And while I cannot control for garment - line chief fixed effects, as each line chief only produces a garment once for the first time, I can control for garment type - line chief fixed effects, to capture interaction effects between lines and classes of garments, such as t-shirts, polos, or pants. Thus, column 4 uses line chief - garment type fixed effects instead of garment fixed effects, which leads to a large increase in the estimated effect of the intervention. This could be indicative of the effect of the

Table 3.5: Main Experimental Results

	(1)		(2)		(3) Non-First Starts only		(4) Non-First Starts only	
VARIABLES	Efficiency	se	Efficiency	se	Efficiency	se	Efficiency	se
Treatment, Day 1	4.096**	(1.76)	6.320***	(0.00)	5.605**	(2.58)	11.049**	(5.37)
Treatment, Day 2	2.483	(3.15)	3.249	(2.59)	3.659	(2.98)	7.850*	(4.77)
Treatment, Day 3	1.252	(2.45)	1.097	(1.98)	2.151	(2.85)	1.681	(2.08)
Constant	42.354***	(0.00)	48.193***	(0.00)	72.414***	(0.00)	37.688	(24.88)
Observations	4,094		4,094		2,743		2,338	
R^2	0.436		0.729		0.748		0.709	
Rank-Grmt.Day FE	YES		YES		YES		YES	
LC-Grmt.Day FE	YES		YES		YES			
Garment FE			YES		YES			
LC-Grmt.Day-G.Type FE							YES	
Week-Fact. FE	YES		YES		YES		YES	

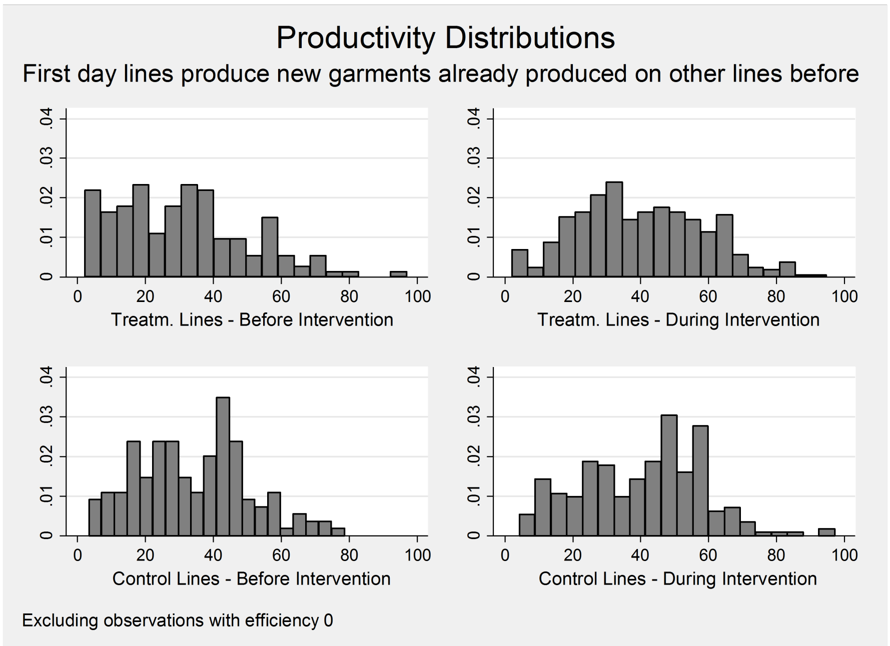
Notes: Table shows the results from regressing daily line efficiency from the first three days a sewing line produces a new garment it has not produced before (first three garment-days) on a dummy that the garment start should have been treated, interacted with garment-day (Treatment, Day 1), and on fixed effects for the number of lines that already produced the garment before, interacted with garment-day, fixed effects for the line chief who started the garment on his line, interacted with garment day, and fixed effects on the week-factory level. Column 2 and 3 add garment fixed effects. Column 4 interacts line-chief - garment-day fixed effects with fixed effects for the type of the garment (t-shirt, polo, pant,). All results estimated using reweighted data according to DiNardo et al. (1996) to balance first day efficiency in pre-treatment period (1st April-20th May 2014). Standard errors obtained by wild-bootstrap, clustered on 17 units of randomization. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

communication intervention being larger among those garments which the line chief was usually worse at producing. Furthermore, using this specification, the effect of the intervention also becomes marginally significant on the second day of production of the garment.

The hypothesis that the effect of the intervention could have had a stronger effect among garment types which the specific line chief would struggle more with also fits with evidence from Figure 3.7, which shows distributions of line productivities on the first day they produce garments that were already produced on other lines before, for four different cases: treatment lines before (Jan-May 2014) and during (Jun-Sep 2014) the implementation of the experiment, and control lines at the same times. The increase in first day productivity of treatment lines during the implementation of the intervention seems to be driven by a strong reduction of the left tail of the productivity distribution. The number of starts with very low productivity is greatly reduced, which is indicative of the individual treatments being enacted specifically when very low productivity could have been expected. This also fits with the fact that less treatments were reported in the logbooks than garments starts selected for treatment were shown in the production data. And while the production engineers said that the logbooks underreport the number of actually conducted treatments, they also explained that in cases in which a line chief could be expected to start the garment without any problems, as he or she had already produced very similar garments before, no treatment was done as no effect of the treatment was expected.

As an additional check on whether this treatment effect could be caused spuriously by a change in the characteristics of garments on treatment lines with the onset of the intervention, Figure 3.8 replicates the distribution graphs of Figure 3.7, however using the SMV of the garments produced on treatment and control lines before and after the start of the garment instead of first day efficiency. Given that the SMV of a garment captures the required labour input to produce one piece of it, the SMV is a proxy for the complexity of the garment, and is highly negatively correlated with efficiency in the overall data. However, there seems to be no shift in the distribution of SMVs of the garments on treatment lines with the onset of the intervention. Average SMV actually slightly increases. There is also no differential pattern visible on control lines. This speaks against the treatment effects being caused spuriously by treatment or control line shifting to a different composition of the garments they produce with the onset of the treatment.

Figure 3.7: First Day Productivity Distribution before and during Intervention

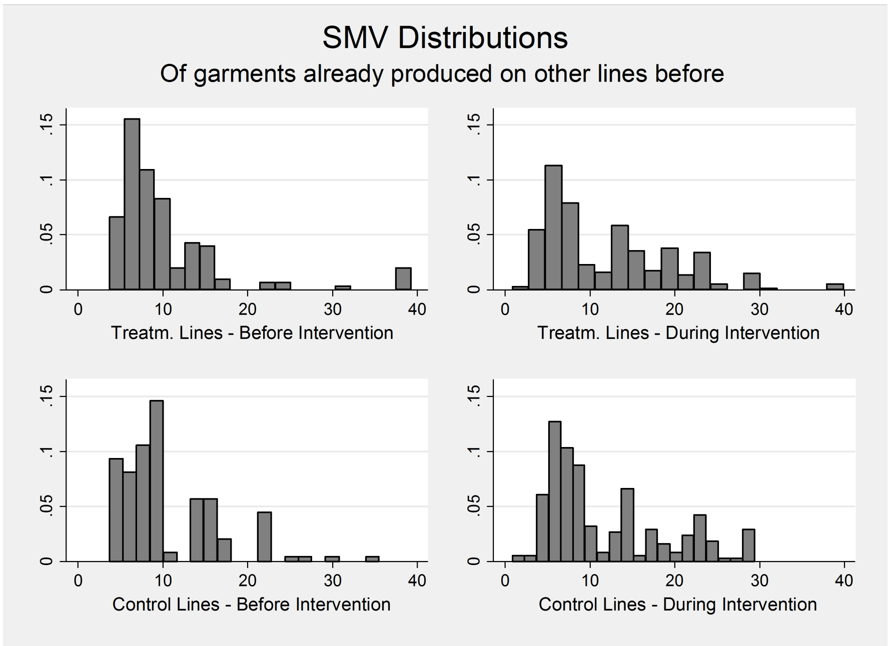


Graph shows distribution of productivity on first day sewing lines produce new garments already produced on some other line in the factory before, on treatment floors, before and during implementation of intervention (top row), and on control floors, before and during the implementation (bottom row).

3.5 Randomized Help Provision and Social Ties

The surveys that I conducted among all line chiefs in the three factories in the sample contained questions on social ties they had to other line chiefs in the same factory. This allows to study whether the effect of the treatment is affected by the presence of social ties to the line chief who provided the briefing. Social ties have been shown to play important roles within firms, such as matching firms with workers (Granovetter [1973, 1995], or Heath [2015] in the case of Bangladeshi garment factories), or to affect effort choice of workers (Bandiera et al. [2010]). Furthermore, they have also been shown to play important roles outside organization in learning about new technologies in small scale agricultural settings (Foster and Rosenzweig [1995]; Munshi [2004]; Bandiera and Rasul [2006]; Conley and Udry [2010]) or about new microfinance services (Banerjee et al. [2013]; Cai et al. [2015]). However, due to the

Figure 3.8: SMV Distribution of Garments before and during Intervention



Distribution of SMV of the garments already produced on some other line in the factory before, on treatment floors, before and during implementation of intervention (top row), and on control floors, before and during the implementation (bottom row).

lack of documentation of many instances in which help should have been provided, it is often not clear which line chief was sent for the briefing. Furthermore, to the extent that the factory management is aware of social ties among line chiefs, it could have sent socially connected line chiefs to provide the briefings in selected instances in which it expected the line chief to profit either particularly much or little from the briefing, which would bias the estimate of the interaction effect. To address this problem, I exploit the instructions that were given to the factories which said that the most 'senior' line chief in the factory who already produced the garment should be the one that is sent to provide the briefings.[18] Therefore, I interact treatment

[18] This specific instruction was given to minimize possible resistance among line chiefs against the intervention, in cases in which less senior line chiefs were sent to more senior line chiefs. In these cases, help provision might have not been accepted by the line chief who was supposed to be briefed. Given that only randomly selected line chiefs receive briefings by these senior line chiefs, the intervention essentially estimates the treatment effect of receiving help from the most senior line chief who already produced the same garment. This estimate might be the most useful for policy implications, as other factories implementing such a management intervention could likely adapt the same policy, of always sending the most senior worker who has experience with certain processes to instruct co-workers on them.

with whether the line chief receiving the treatment reported social ties to the most senior line chief who already produced the garment, which was the case in 59 out of the 377 garment starts which should have been treated.[19]. I measure seniority by the time a line chief already worked as line chief in the factory. To the extent that it was not the most senior line chief according to this measure who was sent to provide the treatment, or that no treatment occurred, the estimated effect can be interpreted as an intention to treat (ITT) effect of the interaction.

Column 1 of Table 3.6 shows the results of this interaction, replicating column 1 from Table 3.5, but adding two further dummy variables, each interacted with garment-day. The first ('Connected, Day n') indicates that the line chief starting the garment shares social ties with the most senior line chief who has so far produced the garment, and the second is an interaction of this variable with whether the garment start should have been treated. In this specification, a positive but insignificant effect can be seen on the first day a line produces a garment, while the effect turns negative, and still insignificant on the second day.

In column 2 the usual line-chief - garment-day fixed effects are interacted with garment type fixed effects, as was already done in column 4 of Table 3.5. Now, the effect of the interaction becomes very large, 17 efficiency units, or 84% of the standard deviations of productivity on the first day a line produces a new garment that has been produced on another line before. However, this large effect comes against the backdrop of a large and negative (but insignificant) effect of being connected to the most senior line chief in general. This could point towards the effect being driven by a few influential observations, and it indeed seems that two observations with unusual high efficiency values of over 150 points have an over proportional effect on the results. I thus drop all 3 observations with efficiency values of more than 150 points.[20] The size of the interaction term now drops to 11 efficiency points (and its p-value to 0.051), while the general effect of being connected to the most senior person who produced the garment so far drops to 6 efficiency points, remaining insignificant.

Thus, there is some cautious evidence that the effect of the treatment was

[19]The network data I use is directed, in the sense that line chief A is considered socially connected to line chief B if and only if line chief A reported a link on one of the six dimensions asked to line chief B, regardless of whether line chief B reported a connection to line chief A or not.

[20]This is done after the data had already be cleaned of clear outliers with efficiency values of more than 200 points.

Table 3.6: Communication Intervention and Social Ties

	(1)		(2)		(3)	
					Effic. < 150	
VARIABLES	Efficiency	se	Efficiency	se	Efficiency	se
Treatment, Day 1	3.278**	(1.67)	4.051**	(1.82)	4.510**	(1.91)
Treatment, Day 2	3.452*	(2.04)	3.880*	(1.93)	4.140*	(2.09)
Treatment, Day 3	1.009	(1.46)	-0.139	(2.34)	0.180	(2.39)
Connected, Day 1	-1.972	(3.98)	-9.169	(5.33)	-6.658	(3.89)
Connected, Day 2	-1.765	(1.67)	-2.161	(2.93)	-2.122	(2.93)
Connected, Day 3	-3.507	(3.91)	-2.457	(3.24)	-2.355	(3.27)
Treatm. x Conn., Day 1	4.170	(4.85)	17.799**	(7.24)	11.855*	(5.51)
Treatm. x Conn., Day 2	-4.245	(7.47)	-3.132	(10.01)	-3.198	(9.93)
Treatm. x Conn., Day 3	2.457	(10.22)	1.569	(8.86)	1.469	(8.92)
Constant	42.169***	(0.00)	45.652***	(4.44)	39.500***	(5.05)
Observations	4,094		3,381		3,379	
R^2	0.438		0.657		0.667	
Rank-Grmt.Day FE	YES		YES		YES	
LC-Grmt.Day FE	YES					
LC-Grmt.Day-Grmt.Type FE			YES		YES	
Week-Fact. FE	YES		YES		YES	

Notes: Table shows the results from regressing daily line efficiency from the first three days a sewing line produces a new garment on a dummy that the garment start should have been treated, interacted with garment-day ('Treatment, Day n'), on a dummy that the line chief of the line reports social ties to the most senior line chief who already produced the garment, interacted with garment-day ('Connected, Day n'), and on an interaction term of the two for each garment-day ('Treatm. x Conn., Day n'). It controls for fixed effects for the number of lines that already produced garment interacted with garment-day, for the line chief who started the garment interacted with garment day, and on the week-factory level. Column 2 and 3 interact line-chief garment-day fixed effects with fixed effects for the type of the garment (t-shirt, polo, pant,). All results estimated using reweighted data according to DiNardo et al. (1996) to balance first day efficiency in pre-treatment period. Standard errors obtained by wild-bootstrap, clustered on 17 units of randomization: *** p<0.01, ** p<0.05, * p<0.1

stronger when the briefing was done by a line chief with whom the line chief receiving the briefing shared social ties, if we condition the regression on line-chief - garment type fixed effects. This is in line with findings from the previous section, that also the overall treatment effect was stronger when conditioned on these fixed effects. In fact, once an interaction term with social ties is included in the regression, the size of the general effect of the treatment remains close to the ones estimated without line-chief - garment type fixed effects. The increase in the overall effect we saw in Table 3.5 when including these fixed effects seems to be driven by those instances in which the briefing was done by a socially connected supervisor.

3.6 Conclusion

This paper presented evidence from a randomized communication intervention in three Bangladeshi garment factories, with the aim of reducing the costs of sharing knowledge about the production processes of garments which are produced on different sewing lines in the same factory. In the intervention, supervisors of randomly selected sewing lines receive a briefing whenever they start producing a new garment which has already been produced by any other line in the factory before, by the most senior supervisor who already produced the garment on his or her line. I show that productivity of those lines was increased whenever they were selected to receive a treatment. The increase in productivity was mainly visible on the first day a line produced a new garment, while in some specifications also on the second. Thus it is visible during the steepest part of the learning curve through which lines go when starting to produce a new garment. The effect was driven by a reduction of the number of starts of garments with extremely low productivity. This points towards the treatment having been more effective at those garment starts at which productivity would have otherwise been very low.

To obtain an idea of the overall effect of this intervention on factory productivity and profits, I use the baseline estimate from column 1, Table 3.5, which shows a significant increase of productivity of about 4.1 efficiency points on the first day a line produces a new garment. Sewing lines on average switch to a new garment every 10 days, and at roughly every second of these starts, another line has already produced the garment before. As average daily productivity in the three factories is 47.4 efficiency units, a very basic back-of-the envelope calculation shows that output was increased by $4.1/(2*10*47.4) \approx 0.43\%$. Anecdotal evidence shows that labour costs make up around 12% of revenue on average in these factories, while the profit margin is about 6%. If we assume that the intervention would save 0.43% of labour costs, this would translate into an increase in profits of 0.86%. On the other hand, the pure monetary costs of the intervention are very low. The hourly wage of a line chief in the factories is about US \$0.50, therefore the wage cost of a half-hour briefing is about \$0.25. In the largest of the three factories with more than 180 sewing lines, if every garment start at which the garment was already produced on another line before were treated, roughly 3000 such briefings would have to be conducted per year, yielding a yearly monetary cost of the intervention of \$750. I do not have information on the revenues of the firms, but local business newspaper report that factories of this size generate revenue in excess of \$10 Mio per year, which, using the commonly referenced margin of 6%, would yield profits of \$600,000. A 0.85% increase would thus imply an increase in profits of ca. \$5,000, or a return on the intervention of more than 700%.

Chapter 4

Social Network Formation in the Workplace: Evidence from Four Bangladeshi Garment Factories

4.1 Introduction

Social networks have been shown of importance in areas as diverse as matching supply and demand in markets, diffusion of new technologies, in driving peer effects, or facilitating contract enforcement.[1]

This growing interest in social networks has fostered the collection of original network data, especially in development economics. Recent studies that do so include Banerjee et al. [2013], who collected network data from 75 villages in Southern India, Beaman et al. [2015], who did so from 200 villages in Malawi, Alatas et al. [2012], who collected network data from 640 hamlets in Indonesia, or Cai et al. [2015], who did so in 185 Chinese villages. These novel network data from rural communities in developing countries add to an existing stock of network data from developed countries, such as the AddHealth dataset, which contains network data among students at 84 US high schools, and which has been extensively used in empirical network research (e.g. Clark and Loheac [2007]; Calvo-Armengol et al. [2009]; Currarini et al. [2009], or Mele [2013]), or the social network data among Harvard students, first used by Leider et al. [2009]. This paper introduces and discusses a novel network data set

[1]Section two of this chapter contains a broad discussion of the vast literature on social networks in economics and social sciences.

collected among sewing line supervisors ('line chiefs') in four Bangladeshi garment factories in 2013/14, and compares characteristics of these networks against those from other networks used in the literature. To the best of my knowledge, this is the first network dataset from workers within large firms in a development context. It is similar in size and focus to the one collected by Bandiera et al. [2010] from workers at a UK soft fruit producer, and it is used in Chapter 2 of this thesis, to estimate whether communication of production knowledge between co-workers within firms has a larger effect on productivity if workers share social ties.[2] This chapter applies several common network formation models to this data to understand which, if any, of these models is best able to explain the processes that led to the formation of these networks at the garment factories.

While the main contribution of this chapter is the introduction and description of the new line chief network dataset, the availability of network data from a novel background offers contributions to a few not yet settled discussions in network research, by adding empirical evidence in favour or against certain hypotheses in the field. First, it shows that in a network in which nodes are grouped into clearly demarked groups (sewing floors), a block-random graph model with just two estimated probabilities, one for within group tie formation and one for cross-group tie formation, does a good job at capturing both the empirical density and clustering levels of the network, something that many other network formation models have struggled to do in other contexts. It thus contributes to a discussion on whether the commonly observed high clustering levels of empirical networks are mainly due to homophily or network externalities (Graham [2015a]). Second, the negative correlation of a line chief's in-degree, but not out-degree, with subsequent turnover adds to a literature in organizational studies which has found conflicting evidence on which measure of network positions matters for predicting turnover in different datasets (Mossholder et al. [2005]; Feeley et al. [2008]). And third, the data shows that newly arriving line chiefs tend to form ties to existing line chiefs with higher lagged in-degree values from the time before the arrival of the new line chiefs, confirming a main assumption of preferential attachment random graph models (even though this model's predictions fit poorly with the characteristics of the line chief network). The availability of lagged degree allows to better identify the effect of degree on subsequent line formation, which is more problematic if only one cross-section of network data is available.

[2]More precisely, a subset of the network data is used in Chapter 2, as only three out of the four factories in the sample also participated in the communication intervention analysed in Chapter 2.

The network data was collected from sewing line supervisors in four garment factories in and around the Bangladeshi capital Dhaka. These factories, locally owned and managed, but producing all of their output for export to international garment retailers, are organized into several departments – knitting, dyeing, cutting, sewing, and finishing. The sewing departments are the most labour intensive ones, employing 50-70% of all workers in the sample factories, and are organized into parallel, independent sewing lines of 20-80 workers, each designed such that the whole sewing process of one garment can be done on one line.[3] Please refer to Chapter three of this thesis for more background information on the factories in the sample.

Each sewing line is headed by a supervisor, typically called 'line chief'. It is these line chiefs among which the network data was collected. All line chiefs at the four factories in the sample were asked to report social ties to all other line chiefs in the same factory on six dimensions:[4] kinship, which other line chiefs one already knew before working at the factory, with whom one had already worked together at another factory before, whom one had already visited at home, with whom one spends lunch breaks together, and with whom one generally is befriended.

The surveys were administered twice, in 2013 and 2014, at Factory 1 and 2, and once in 2014 at factory 3 and 4, as these two factories only joined the project in early 2014. Thus, the dataset contains six networks on the factory-year level. At Factory 1, 67% of all line chiefs surveyed in 2014 were already surveyed in 2013, while this share was 47% in Factory 2, implying rather high turnover rates among line chiefs in these factories. In total 227 surveys of 185 distinct line chiefs across the four factories were conducted, implying that 42 line chiefs were interviewed twice, once in 2013 and once in 2014. Table 4.1 below summarizes key demographic characteristics of the surveyed line chiefs for each factory-year. Only two out of the 185 line chiefs interviewed are female (with one female line chief interviewed twice), reflecting a very strong gender imbalance at supervisory levels at the factories, which can

[3]The number of workers per line is much more homogeneous within the sample factories. In fact, within factories, lines are kept more or less interchangeable by the factory management, to maximize flexibility in a production environment characterized by many shocks, such as power failures, supply and delivery chain failures (often caused by the frequent political general strikes, called 'hartals'), and low predictability of the orders which will be placed by buyers. Variation across factories in average number of workers per line stems from specialization of factories into garment types (shirts, pants, jackets,) or fabric types (knit, woven).

[4]The four sample factories are located 1-3 hours travel by car from each other. Thus, it is unlikely that the surveyed line chiefs might have social ties to line chiefs at another of the four factories in the sample, and no attempt was made to elicit such ties across factories in the surveys.

Table 4.1: Line Chief Characteristics

	Fact. 1 2013	Fact. 1 2014	Fact. 2 2013	Fact. 2 2014	Fact. 3 2014	Fact. 4 2014
Age	28.8	30.0	30.4	30.0	32.2	30.8
Years in factory	3.9	4.3	5.9	5.0	6.2	5.2
Years as Line Chief	2.0	2.4	1.9	1.8	3.5	2.3
Promoted Internally	56%	58%	81%	73%	52%	77%
Education Code	14.9	15.0	14.9	14.7	15.7	15.1
N	57	53	16	15	60	26
N Female Line Chiefs	1	1	0	1	0	0

Notes: All information from survey of all line chiefs in factory. Promoted internally is percentage of current line chiefs who worked on lower position in same factory before and were subsequently promoted to line chief.

be found throughout the Bangladeshi garment industry.[5] Otherwise, average line chief characteristics are homogeneous across the factories, particularly age. Only at Factory 3 are line chiefs slightly older, and have worked already slightly longer at the current factory, as well as on their current position as line chiefs. They also report on average slightly higher educational attainment.[6]

This paper proceeds as follows. Section two will survey the literature on social networks. Section three will describe in more detail the network data collected from the line chiefs at the four factories. Section four will apply basic random graph network formation models to the data, while section five will do so with models which allow for richer heterogeneity in network nodes. Section six will study actual network formation directly from the data, exploiting the fact that from Factory 1 and 2, two network surveys at different points in time are available, allowing us to directly observe how line chiefs who arrive new between these two surveys form ties

[5]Notwithstanding this, ca. 80% of the total sewing labour force in the four sample factories is female. Workers typically start working in the garment factories at the age of 18 (child labour prohibition these days enforced through foreign buyers), and drop out again at the age of 25-30, unless promoted to supervisory, mechanical, or quality inspector positions. However, almost only male workers are promoted to these positions. Chapter 2 of this dissertation compares the effectiveness of female and male supervisors from a multi-angle perspective, and discusses possible reasons for the very low promotion rate of women to supervisory positions.

[6]Differences in age, seniority and education from Factory 3 to other factories are statistically significant on 5% level. Education code is an IPA Bangladesh (the survey firm) specific code for educational attainment, with 14 implying 10 years of schooling, 15 implying secondary schooling certificate (SSC), and 16 implying 11 years of schooling.

with the existing line chiefs and with other new ones. Section seven will study the effect of network position on turnover between the two survey rounds, while section eight will conclude.

4.2 Literature

The literature on social networks is vast, and this chapter can only provide a tentative overview. The research on social networks originated from sociology and political science . However, at least since the 1990s, interest into social networks has rapidly grown among economists as well. The literature on networks can broadly be separated into two classes. The first takes social networks as given, and studies its effects on people's behaviour, knowledge, and attitudes, while the second studies the formation of networks. However, the formation of networks is often endogenous to its effects. Therefore, this separation is not always desirable. Only recently, though, did studies emerge which analyse the formation and effects of networks simultaneously (e.g. Goldsmith-Pinkham and Imbens [2013] on peer effects in academic achievement among high-school students).

4.2.1 Literature on Effects of Social Networks

The first class of the literature on the effects of given networks is the older and larger one, spanning social sciences such as sociology, political science, anthropology, economics or criminology. Lazarsfeld et al. [1944] is often cited as a starting point of research on the effects of social networks, when they show how US voters rely on the opinion of other people in their social network, and especially on opinion leaders among them, to make decisions on whom to vote. Lazarsfeld and Merton [1954] coined the term 'homophily', to describe the often observed preference of persons to form social ties to other persons who are similar to them. Myers and Schultz [1951] and Granovetter [1973, 1995] showed that large shares of jobs are found through social networks of workers, starting a large literature on the interplay between social networks and labour markets. Ioannides and Loury [2004] provide a survey of this literature up to this point. More recently, one focus of the literature on social ties and finding a job has been on immigrant communities. Munshi [2003] uses rainfall as instrument for migration waves of Mexicans to the US, and shows that when larger numbers of Mexicans have recently arrived in the US, the chances of finding a job for newly arriving immigrants increase. Beaman [2012] finds similar effects using exogenous resettlements of refugees from different nationalities among US cities.

The role of social ties in bringing together supply and demand has also been documented in markets other than the labour market. In an early paper about Moroccan bazaar traders, Geertz [1978] argues that focusing attention on a narrow subset of possible trade partners in the market to whom one is socially connected is profit maximizing. Sampling offers of supply and demand for a given good from all traders in the market on any given day does not allow to collect the necessary background information on all the traders to prevent being shortchanged. Uzzi [1996] documents the role of social ties among producers and traders in the New York garment industry. Social ties allow for the transfer of valuable information between traders and producers and facilitate joint problem solving in case of unforeseen difficulties. He goes on to show that traders that conduct more trade within steady business relationships are less likely to exit business. Weisbuch et al. [1996, 2000] are case studies about the fish market of Marseille, France, showing how many buyers only buy from a small subset of sellers in the market. This tendency is more pronounced among buyers which buy more frequently and larger amounts. Finally, McMillan and Woodruff [1999] find that Vietnamese entrepreneurs whose relationship was initiated through business associations or common third business partners grant more trade credit to each other, as these social ties can be leveraged against defaulting on the credit.

Another widely studied effect of social ties is the one on technology diffusion. The literature can be traced back to Coleman et al. [1966], who showed that physicians who had more social ties were faster at starting to prescribe novel drugs. They argued that this is caused by the spread of information on these drugs through social networks. Using a more careful research design, by looking at the effect of 'opinion leaders' among physicians during times of heightened uncertainty due to changes in prescription guidelines, Bhatia et al. [2006] confirm the effect of social ties among physicians on drug adoption. Over the last decades, the effect of social ties on technology adoption has been studied especially intensely in the context of small scale agriculture in developing countries. Foster and Rosenzweig [1995] and Munshi [2004] in the context of India, Bandiera and Rasul [2006] in Mozambique, and Conley and Udry [2010] in Ghana all documented how the adoption of improved seeds or more profitable cash crops spread along social ties among farmers. Similarly, in a recent paper, Banerjee et al. [2013] inform exogenously selected leader in Indian villages about newly available microfinance services, and show how knowledge and take-up of these services spread along social ties within the villages. The take-up decision in this case was mainly driven by knowledge that spread about the product, not by the number of other people in ones social circle who also took up the product. Thus the effect of social ties on adoption seems to be primarily a knowledge diffusion effect, and not so much a peer effect. Similarly, Cai et al. [2015] show that adoption of rain-indexed agricultural insurance spreads along social networks among Chinese villagers. They argue that as in the case of microfinance take-up in the Indian villages, the effect was driven by diffusion of knowledge about the product, and not by peer effects. Similar research in developed countries has shown how the decision on whether and which health insurance plan (Sorensen [2006]) or retirement plan (Duflo and Saez [2003]) to pick is strongly correlated among colleagues working in the same departments within organizations. Duflo and Saez [2003] provided a small monetary incentive to random workers from a large organization to attend retirement plan information events, and find that non-incentivized worker from the departments of the incentivized workers have a significantly higher propensity to attend the events as well, and to obtain retirement plans, compared to non-incentivized workers from departments at which no worker was incentivized. Bertrand et al. [2000] showed how social networks affect dependency on welfare, while Hong et al. [2004] show their effect on participation in the stock market.

A further strand of literature has emphasized the role that kinship and neighbourhood networks play in insuring against negative shocks in developing countries (De Weerdt and Fafchamps [2011]; Fafchamps and Gubert [2007a,b]; De Weerdt and Dercon [2006]). Two main findings of this literature are that these networks are not necessarily formed optimally, in that they do not link heterogeneous members which likely face uncorrelated shocks. Furthermore, transfers to insure against negative shocks are not necessarily based on reciprocity but exhibit significant net-transfers over time between network members, pointing towards social norms and altruism as driver behind these transfers. In a recent interesting contribution, the ability of community based networks to sometimes shift whole communities into new occupations with better income prospects has also been described (Munshi [2011]). However, community network effects need not be unambiguously positive, as rural community based insurance networks can also discourage people to migrate to cities in search of better paid work (Munshi and Rosenzweig [2013]).

Finally, the work of Kandel [1978] on how peers in social networks of adolescents affect political attitudes and the propensity for drug use spawned a long interest in social sciences on networks and peer effects among adolescents. Recent contributions are for example Clark and Loheac [2007], who confirm peer effects among adolescents on (legal) drug use, showing that these effects are stronger among boys, or Calvo-Armengol et al. [2009], who show that the centrality of the position of a student in his or her peer network affects his or her school performance.

4.2.2 Literature on Network Formation

The second class of research on networks studies the formation of networks. For the most part, this research attempts to infer from the structure of established networks clues on its formation process, by utilizing models of network formation which yield distinct predictions for the resulting networks. Thus, the final network structure allows to some extent to discriminate between different models in terms of the likelihood with which they describe the true process through which the network was formed.

The literature on network formation can be split up into two subcategories, random and strategic network formation. Random network models can be traced back to the seminal paper of Erdos and Renyi [1959], who study the properties of networks in which each link between two nodes of a network is created randomly and independently with a given probability. Even though the model is too stylized to explain many features of real world networks, it has ever since served as the foundation for most of the empirical network formation literature.

Strategic network formation models, which assume that link formation in a network is not a random process but the outcome of utility maximization of two or more nodes, have emerged more recently. Jackson and Wolinsky [1996] are often cited as the starting point of this literature, who introduced widely used equilibrium concepts, such as pairwise stability, to study which networks can emerge in equilibrium given the utility functions of the network members. The focus of this literature is on the theoretical modelling on network formation, and the testing of the predictions mostly utilizes again empirical models founded on random graph theory. Given the focus of this paper on a novel empirical network dataset and the estimation of several empirical network models on this data, the discussion of theoretical models of network formation is largely skipped, and only touched upon insofar as it helps guiding the application of empirical models on the new dataset. See Jackson [2008] for a discussion of basic theoretical strategic network models.

Basic random graph models are fitted to the data using simple network density values, and predictions of the model regarding clustering values and degree distributions can then be compared against their empirical counterparts. The estimation of more complex models which are often used to test the predictions of theoretical network models, and which can incorporate characteristics of individual network nodes, are faced with more difficulties, and the literature is still rapidly evolving. One strand of this literature regresses the existence of links on the node-pair level on observable characteristics of the involved nodes (e.g. Fafchamps and Gubert [2007a,b], or, more recently, Graham [2015b]). However, these studies assume that link formation, conditional on observables, is still independent for each pair of two nodes, thereby ruling out externalities in link formation, which are at the heart of many strategic network models. To estimate network formation allowing for externalities in link formation, researchers have used exponential random graph models (ERGMs). The estimation of these models is technically sophisticated and requires sampling from distributions over networks. Furthermore, the commonly used sampling methods have recently been criticised for being unreliable (see Chandrasekhar [2015] for a recent discussion of ERGMs and other empirical network models). Given both their technical complexity and the recent uncertainty about their estimation, I will not estimate ERGMs in this paper.

A recent contribution to the estimation of network generation models under the presence of link externalities by Chandrasekhar and Jackson [2014] are subgraph models (SUGMs). These models combine the ability of ERGMs to capture arbitrary levels of link externalities with the easy estimateability of node-pair level regressions. They assume that not only pairs of nodes can decide whether or not to form links between them, but groups of nodes of arbitrary size can agree to form specific sub-graphs among them, such as cliques (complete connection among all nodes in the group) or stars (all nodes in the group are connected to one node among them). By assuming a model in which groups of certain sizes can form certain subgraphs, these models allow estimating parameters of network formation models which can reproduce models that exhibit high levels of dependency of link formation. Section five will estimate a simple expositional SUGMs model based on Chandrasekhar [2015], using the line chief network data.

4.3 Description of Line Chief Network Data

This section introduces the line chief network data in more detail. The directed graphs of the six networks on the factory-year level are shown in Appendix E. In the graphs, the different colours of the rectangular nodes indicate line chiefs working on different sewing floors, and the thickness of the links that connect the nodes represents the number of different dimensions on which a connection was reported by one line chief to another. The networks are all directed, which means that the existence of a link from line chief i to line chief j does not imply the existence of a link from j to i. Especially if network data is only available from a sample of the target population, directed links are transformed into undirected links in many empirical studies, by assuming a link in both directions between nodes i and j as long as reported by at least one of the nodes (Banerjee et al. [2013], Chandrasekhar et al. [2014]). The subsequent analysis, however, always works directly with the directed network data, unless explicitly stated otherwise, as some analysis is only possible with, or yields results that are better interpretable, when using undirected network data. In these cases, the network is transformed into an undirected network as described above, by assuming a link in both directions between two line chiefs as long as reported by at least one of them.

In total, 620 directed links have been reported across the 227 line chief surveys, of which 278, or 44.8% were reciprocated by the other line chief. This ratio seems to be broadly in line with other empirical network studies. For example, in the network among Harvard students from Leider et al. [2009], 36.6% of all links were reciprocated, while Feeley et al. [2008] report a reciprocity rate of 30% in a social network among fast-food employees. If we assume a directed link from one line chief to another as long as at least one of the two line chiefs has reported a link in either direction, we would therefore have 962 such directed links, or 481 non-directed links between two line chiefs.

4.3.1 Degrees

As a first description of the networks, Table 4.2 shows the average out-degree, the number of links a line chief reports to other line chiefs, for the six different networks on the factory-year level. It distinguishes between links to line chiefs from all sewing floors in the factory (column 1), and only from the same floor (column 2). On average, five (Factory 2 and 3) to ten (Factory 1 and 4) line chiefs work on

Table 4.2: Average Out-Degree of Line Chiefs

Fact.	Year	(1) Out-Degr.	(2) Out-Degr. Same Fl.	(3) Weighted Out-Degr.	(4) Weighted Out-Degr. Same Fl.,	N	#Flrs
1	2013	2.28	2.07	4.58	3.88	57	6
1	2014	2.81	2.77	4.38	4.28	53	6
2	2013	3.75	2.31	6.81	4.19	16	4
2	2014	4	2.73	6.87	4.67	15	3
3	2014	2.15	1.87	3.72	3.18	60	14
4	2014	3.54	3.23	5.54	4.81	26	3
All		2.73	2.37	4.72	3.97	228	36

Notes: Table shows for each network on the factory-year level average out-degree of a line chief, when considering links to all other line chiefs in the factory (column 1 & 3), and to all line chiefs on the same floor (column 2 & 4). Column 3 & 4 use weighted outdegrees, where weighted links mean that a link reported from one line chief to another is multiplied by the number of dimensions it is reported on (visited home, lunch breaks, friendship,...).

the same sewing floor, which are roughly equally sized within factories (see nodes of same colour in network graphs in Appendix E). Furthermore, while column 1 and 2 show the average number of links reported on at least one dimension, column 3 and 4 show the average number after weighing links by the number of dimensions they are reported on. A link reported on only one dimension has weight 1, a link reported on two dimensions has weight 2, and so on. Essentially, in the weighted links degree measure, a link reported to another line chief on one dimension, e.g. spending lunch-breaks together, is counted as a separate link from one reported to the same line chief on another dimension, e.g. 'being friends'.

On average, line chiefs report links to 2.73 other line chiefs in the factory, of which on average 2.37 are reported to line chiefs from the same floor. This shows that social ties are heavily concentrated within floors. When looking at weighted links, line chiefs report on average 4.72 links, which implies that the average link is reported on 4.72 / 2.73 = 1.73 dimensions, while ties to the same unit are reported on average on 1.67 dimensions. At no factory and year are ties to the same unit reported on average on more dimensions than ties to other units. Thus, while being reported more rarely, conditional on being reported, ties to other units are reported on slightly more dimensions. This could be indicative of ties to line chiefs on other floors being underreported relative to those from the same unit. It could be that line chiefs from the same floors were more salient in the mind of line chiefs, when being asked about ties to fellow line chiefs. Therefore, also relatively weaker ties to line chiefs from the same floor were reported, while ties to line chiefs on other floors had to be of a higher minimum strength to be reported.[7]

Table 4.3 below shows how many ties were reported on average by the line chiefs on the six different dimension, again separately between ties to all line chiefs, and those from the same floor. The number of kinship ties seems negligible, which indicates that the hiring process in these factories is not characterized by current workers referring family member for jobs if the factories look for additional workers (though Heath [2015] argues that many jobs in the Bangladeshi garment industry are filled through factories asking experienced workers to refer acquaintances from their home villages to them, in case positions need to be filled). Also only a minority of workers report ties to co-workers in the factories which they knew already before working at the factory. Most ties are reported on the dimensions of spending lunch breaks together, and simply regarding each other as 'friends'.

The top row of Figure 4.1 below shows the out-degree distribution among all line chiefs in the six networks. The degree distribution is of prime importance in network research, as its shape constrains many higher order network characteristics (Faust [2007]), and as it is an important statistic for fitting random network models to empirical data. The left-hand side graph shows the non-weighted distribution (each link has the same weight, regardless of the number of dimensions the link was reported on), while the right-hand side graph shows the distribution with links weighted by number of dimensions on which they are reported. Both empirical distributions seems to loosely follow a power distribution, which is a common feature of empirical network data (Jackson [2008]). I will return to fitting statistical distributions to the network data in the next section, when I discuss random graph network formation processes in the context of my data. The bottom row of Figure 4.1, on the other hand, shows the distribution of node's in-degree, the number of directed links to the line chief reported by other line chiefs. These distributions seem to be

[7]In the surveys, line chiefs were asked, dimension by dimension, first "to which other line chiefs from the same floor are you connected on dimension X", and then"to which line chiefs from other floors are you connected on this dimension". The interviews were usually conducted in separate rooms just off the sewing floors where the line chiefs worked.

Table 4.3: Reported Social Ties on Individual Dimensions

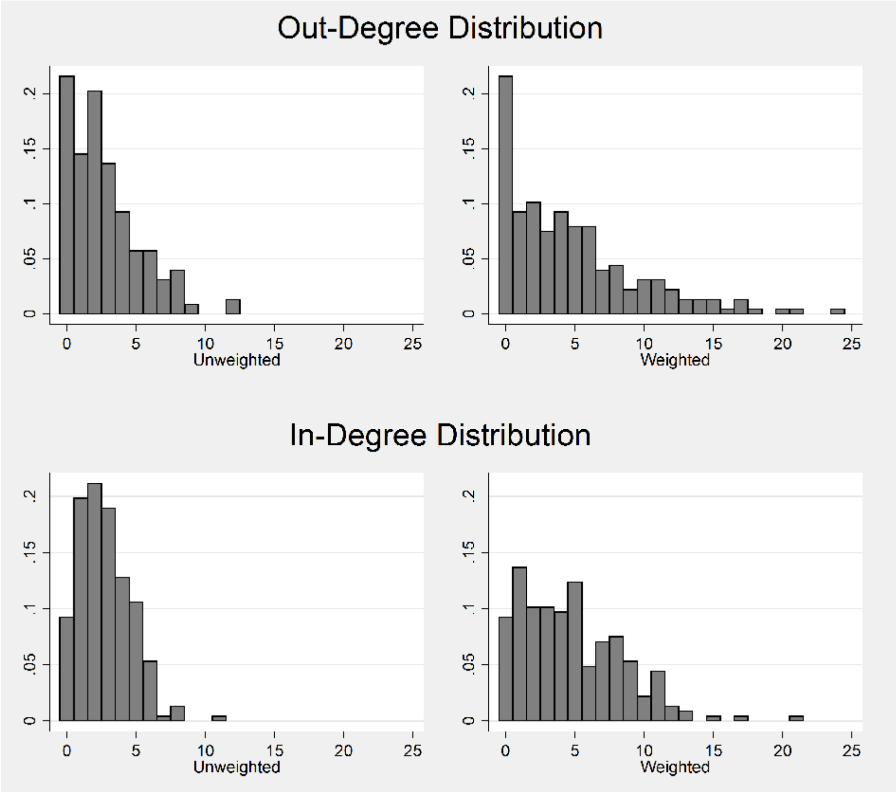
	(1)	(2)
Dimension	Out-Degree	Out-Degree, Same Floor
Kinship	0.02	0.004
Knew Before	0.37	0.26
Worked Together	0.16	0.13
Visited	0.96	0.83
Lunch	1.33	1.15
Friends	1.88	1.58
All	4.72	3.97

Notes: Table shows average number of links reported to other line chiefs on the six individual dimensions asked. Column 2 shows average number of ties reported on these dimensions to other line chiefs located on same sewing floor.

closer to Poisson-distributions, especially the one of unweighted in-degrees, another often observed distribution of degrees in empirical networks. There are also notably less line chiefs with in-degree zero than with out-degree zero. This is also reflected in the lower standard deviation of in-degree, 1.86, compared to 2.53 for out-degree (using unweighted degrees). 44 of the 227 line chiefs do not report any ties to other line chiefs but nevertheless do receive ties from others, while there are only 16 line chiefs who report ties but do not have any incoming ties. Thus, while line chiefs vary more in the propensity to form, or report, links, the links they report seem to be directed more evenly across line chiefs.

Table 4.4 below tests whether observable line chief characteristics correlate with (unweighted) out- and in-degree, by regressing the degrees on line chief observables and factory-year fixed effects. Standard errors are clustered at the line chief level (recall that 42 line chiefs at Factory 1 and 2 were interviewed twice in 2013 and '14). When using out-degree, gender shows a significant effect. The two female line chiefs report more ties. Furthermore, older line chiefs report slightly less ties. Given that degree is a count variable, column 2 repeats column 1 using a Poisson regression model. In this regression, only the effect of gender remains, while the effect of age vanishes. With in-degree, gender flips its sign, now exhibiting a highly negative effect. Much less ties are reported to the two female line chiefs, even though

Figure 4.1: Degree Distribution of Line Chiefs



Graph shows empirical distribution of out-degree of all line chiefs across all factories in top row, unweighted (left graph), and weighted (right graph). Bottow row shows in-degree distribution accross all factories, unweighted and weighted. Weighted links mean that a link reported from one line chief to another is multiplied by the number of dimensions it is reported on (visitied home, lunch breaks, friendship,...).

they themselves report more ties than their male peers. There might be a culturally rooted tendency in this male dominated group of line chiefs to either not form, or not report social ties to female co-workers. Furthermore, seniority as line chief, as represented by years already working as line chief in the factory, has a positive effect on in-degree. When using a Poisson regression model, seniority in the factory, the time one has already worked on any position in the factory, shows a slight additional positive effect. There might be a status effect from being socially connected to senior line chiefs, and thus social ties to them are more eagerly seeked or reported.

Table 4.4: Predictors of Out- and In-degree of Line Chiefs

	(1)	(2)	(3)	(4)
		Poisson		Poisson
VARIABLES	Out-Degree	Out-Degree	In-Degree	In-Degree
Age	-0.062*	-0.024	0.002	-0.001
	(0.037)	(0.014)	(0.023)	(0.008)
Gender	1.142**	0.349**	-2.948***	-14.573***
	(0.518)	(0.173)	(0.342)	(0.735)
Seniority in Fact.	0.057	0.020	0.117	0.042*
	(0.100)	(0.033)	(0.085)	(0.024)
Seniority as LC	0.035	0.016	0.192**	0.063**
	(0.122)	(0.041)	(0.095)	(0.028)
Started this Position	-0.120	-0.035	0.015	0.033
	(0.549)	(0.195)	(0.350)	(0.118)
Education	0.009	0.005	0.029	0.017
	(0.074)	(0.027)	(0.048)	(0.017)
Constant	2.711	1.037	3.894***	14.792***
	(1.816)	(0.660)	(1.066)	(0.836)
Observations	216	216	216	216
R^2	0.084		0.323	
Factory FE	YES	YES	YES	YES

Notes: Table shows results from regressing out- and in-degree of line chiefs on observable line chief characteristics, and factory-year fixed effects. Column 2 and 4 use a Poisson regression model. Robust standard errors clustered on line chief level in parentheses: *** p<0.01, ** p<0.05, * p<0.1

4.3.2 Core-Periphery Structure

The finding that seniority has a positive effect on in-degree points towards the existence of a core-periphery structure of the network, as can be found in many other empirical networks. The concept of core-periphery structures in social networks arguably takes a more prominent role in organizational and business studies, which often found this pattern in social networks within organizations, than in economics. In within-organization networks, membership in the core group often comes with a high social status, derived from seniority or high productivity (Burt [2000]; Fuchs [1995]), which fits with positive correlation of in-degree and seniority in the line chief data. Cummings and Cross [2003] found some evidence that a more pronounced core-periphery network structure was associated with lower performance of work teams.

In its simplest form, a core-periphery structure is a partition of all nodes in a network into two groups (core and periphery), with nodes from both groups preferring to have links with nodes from one group (the core) (e.g. Persitz [2009]). This definition is met by the line chief network data, when splitting the line chiefs in two groups of above and below median seniority, in terms of time a line chief already worked on this position, within each factory-year network. The following Table 4.5 shows which shares of possible links within these two groups, and across these two groups are reported. The highest share of links is found within the group of senior line chiefs. Below median seniority line chiefs are more likely to report ties to above median seniority line chiefs than to other below median seniority line chiefs, with the difference being statistically significant on the 1% level. Senior line chiefs are as likely to report ties to junior line chiefs as junior line chiefs are among themselves, but also more likely to report ties to other senior line chiefs. Thus we have a typical core-periphery structure with a more connected core of more senior line chiefs, and a periphery which is more connected to the core than among itself. This pattern also holds when defining the core as the group of line chiefs with seniority above the 66th percentile or above the 75th percentile, as shown in Table 4.5 as well.

4.3.3 Persistence of Links over time

I can check how persistently ties are reported over time within the subset of line chiefs which were interviewed twice at Factory 1 and 2, in 2013 and 2014. In 2013, 72 directed ties were reported within this subset of 42 line chiefs. In 2014, 37 of these ties were reported again, or 51%. This ratio masks some heterogeneity across the two factories, with the ratio being 39% at Factory 1, and 94% at Factory 2. Note that the lower percentage from Factory 1 does not necessarily imply a low persistence across the survey years. 35 out of the 42 line chiefs interviewed at both years are from Factory 1. This implies that there are 35*(35-1) = 1190 possible directed

Table 4.5: Core-Periphery Pattern

MEDIAN	To below median Sen. LC	To above median Sen. LC
From below median Sen. LC	0.047	0.065
From above median Sen. LC	0.048	0.078
66 PERCENTILE	To below 66 pctl. Sen. LC	To above 66 pctl. Sen. LC
From below 66 pctl. Sen. LC	0.053	0.069
From above 66 pctl. Sen. LC	0.047	0.085
75 PERCENTILE	To below 75 pctl. Sen. LC	To above 75 pctl. Sen. LC
From below 75 pctl. Sen. LC	0.053	0.070
From above 75 pctl. Sen. LC	0.052	0.104

Notes: Table shows likelihood of a line chief below or above median seniority to report a link to another below or above median seniority line chief. Second and third panel of table show numbers when considering line chiefs above and below 66'th and 75'th percentile of seniority. Seniority is measured in years line chief already works in factory as line chief.

links within these 35 line chiefs from Factory 1 per survey round. In the second survey round from 2014, these line chiefs from Factory 1 reported 70 directed links among themselves. Had there been no persistence across the two survey rounds in terms of reported links, the probability that a link that had already been reported in 2013 would be reported again in 2014 would have been the same as if the 70 links would have been allocated randomly (without replacement) among the 1190 possible links. Thus any given link would have been reported with probability 0.061. However, the fact that links that were reported in 2013 were reported again in 2014 with probability 0.393 shows that there is indeed significant persistence in reported links over time. I am not aware of other studies which attempted to measure the persistence of reported social networks, against which the results from this network can be compared against.[8]

[8] Burt [2000] reports high decay in business relationships, defined by bankers doing business deals among each other, with 90% of relationships gone after four years.

4.3.4 Density, Diameter and Component Structure

The density of a network is simply the share of all possible links in the network that are reported. Define $g_{ij} \in \{1, 0\}$ as the element on the ith row and jth column of the adjency matrix of a network g, taking value 1 if a link exists from node i to node j, and zero otherwise. For directed networks, the density of a network g is thus defined as:

$$density(g) = \frac{\sum_{j \neq i} g_{ij}}{n(n-1)} \tag{4.1}$$

Most empirical networks exhibit low density, or 'sparsity' (Chandrasekhar [2015]). More precisely, a 'sparse' network is defined as a network whose density goes to zero if its size n, the number of nodes it contains, goes to infinity. Networks are sparse as long as the average degree of a node increases less than proportionally as n grows. Most empirical networks show a pattern in which average degree stays constant as n goes to infinity, which could be due to nodes having a capacity to maintain only a limited number of ties. Therefore, these networks are sparse.

The density of the line chief network is on average 0.059 across the six sub-networks, varying between 0.036 at Factory 3 and 0.286 at Factory 2 in 2014, as shown in column 1 of Table 4.6. Interestingly, the rank of the six sub-networks when it comes to their density fits exactly with their inverse rank when it comes to their size; the smallest network (Factory 2, 2014) has the highest density, while the largest network (Factory 3) has the lowest. Thus, also the network generation process underlying these networks is likely to be one in which degrees increase less than proportionally with network size, therefore generating sparse networks.

While many real world networks exhibit low density, they often have at the same time a low diameter and low average path lengths. Define as path a connection between two nodes along other nodes in the network which only follows existing links, and which does not uses the same link twice. Define the distance between two nodes as the number of links one has to follow on the shortest path between the two nodes. Diameter refers to the longest distance between any two nodes in the network. Diameter thus refers to the maximum number of links one has to follow in order to reach any node in the network from any other node. Average path length, on the other hand, refers to the average shortest path lengths between all pairs of nodes in the network. The observed prevalence of low average path lengths or diameters in networks with low density has been dubbed the 'small-world' phenomenon

(Milgram [1967]).

Diameter and average path length can easily be calculated also for directed networks, whereby the directed distance between two nodes now is the smallest number of links one has to follow in their defined direction to reach one node from the other. However, diameter and average path lengths can only be calculated for connected networks, in which each node can be reached from any other node by following a series of links. The directed diameter for directed networks, accordingly, can only be calculated for 'strongly' connected networks, in which each node can be reached from any other node following links in the direction they are defined on. As can be seen from the graphs in Appendix E, only the two networks from Factory 2 are connected, and no network is strongly connected. In case of non-connected networks, the diameter is sometimes set to infinity. Alternatively, one could report the diameter for the largest component of the network, which is the largest sub-set of nodes in the network which are (strongly) connected. This is especially common if the largest component is a 'giant component'. Giant components are defined as components which include more than $n^{2/3}$ nodes of the network, and are the only component in the network to do so (Jackson [2008]). Given this definition, those four networks that are not connected do have giant components, at least for the undirected network, as the size of the largest undirected component (column 5, Table 4.6) is larger than $n^{2/3}$ (column 7). However only three of the six networks have a directed giant component (column 4 vs column 7). Columns 2 and 3 of Table 4.6 show the directed and undirected diameter of the largest component of the six sub-networks on the factory-year level, which is relatively small compared to component size, at least for the undirected networks.[9]

4.3.5 Clustering

Many empirical networks exhibit clustering, the existence of sub-groups of nodes which have more links among them than would have been expected if all links in the network were created randomly and independently. A common way of measuring clustering in a network is by the overall clustering coefficient. It measures the share among all instances in which a node i is connected to two other nodes j and k, in which j is also connected to k. More formally, for undirected network data, let $g_{ij} \in \{1, 0\}$ denote again whether there exists a link between node i and j. Then

[9]Note that the directed diameter can be shorter than the undirected diameter, when the largest strongly connected component is smaller than the largest connected one for the undirected network.

Table 4.6: Density, Diameter and Largest Components

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fact.	Year	Density	Dir. Diam.	Undir. Diam.	Largest Dir. Comp.	Largest Und. Comp.	N	$N^{2/3}$
1	'13	0.040	7	8	11	47	57	14.8
1	'14	0.054	7	7	13	29	53	14.1
2	'13	0.250	4	4	11	16	16	6.4
2	'14	0.286	4	3	13	15	15	6.08
3	'14	0.036	7	10	9	45	60	15.3
4	'14	0.142	7	5	18	21	26	8.8
All		0.058					220	

Notes: Table shows for each network on the factory-year level its density, directed and undirected diameter (col. 2 & 3), and size of the largest connected and strictly connected component (col. 5 & 4).

the overall clustering coefficient can be expressed as:[10]

$$C^U = \frac{\sum_{i;j\neq i;k\neq i;j\neq k} g_{ij}g_{ik}g_{jk}}{\sum_{i;j\neq i;k\neq i} g_{ij}g_{ik}} \tag{4.2}$$

Column 1 of Table 4.7 below shows the clustering coefficient for the undirected versions of the six networks. If all links in a network would have been created randomly and independently between any two nodes i and j with probability p, the clustering coefficient itself would also be p. In the networks from the garment factories, 481 out of 5,294 possible undirected links were reported. Had the network been created by such a random process, the ratio $481/5,294 = 0.091$ would be an estimate for probability p, and therefore also for the clustering coefficient. However, as can be seen from column 1 of Table 4.7, the empirical clustering coefficients from the six networks are much higher, ranging between 0.40 and 0.73, thus indicating high levels of clustering. They are also higher, but not out of range, compared to other networks commonly used in the literature (compare the clustering coefficients for various networks shown in Table 1 of Chandrasekhar (2015), which range from 0.17

[10]The overall clustering coefficient can be distinguished from what is sometimes referred to as the average clustering coefficient, which is basically a node level clustering coefficient, with the average taken over all nodes to obtain the final coefficient. For each node i, among all pairs of nodes j and k that are both connected to node i, the share at which j and k are connected is taken to provide a node-level clustering coefficient, which can then be averaged across. This value by construction correlates strongly with the overall correlation coefficient, though the average clustering coefficient gives a higher weight to nodes with low degree.

Table 4.7: Clustering Coefficients

Factory	Year	(1) Clustering Coef. Undirected	(2) Clustering Coef. Directed
1	2013	0.40	0.34
1	2014	0.64	0.56
2	2013	0.52	0.44
2	2014	0.56	0.45
3	2014	0.44	0.38
4	2014	0.73	0.56
All		0.52	0.46

Notes: Table shows for each network on the factory-year level the undirected clustering coefficient as defined in equation 4.2, and the directed clustering coefficient as defined in equation 4.2.

to 0.41). Note that the much higher propensity in the line chief network data to report ties to line chiefs from the same floor, as opposed to line chiefs from other floors, naturally induces a higher level of clustering.

A variation of the clustering coefficient for directed networks such as the line chief network is the share among all cases in which a node i has a link to node j, and node j a link to node k, in which i also has a directed link to node k. More formally, this coefficient can be expressed as a slight variation of equation 4.2:

$$C^U = \frac{\sum_{i;j\neq i;k\neq i;j\neq k} g_{ij}g_{jk}g_{ik}}{\sum_{i;j\neq i;k\neq i} g_{ij}g_{jk}} \tag{4.3}$$

Column 2 of Table 4.7 shows this directed clustering coefficient for the six networks, and the average coefficient across the networks. These coefficients are somewhat lower than the undirected ones, but are still much higher as what would have been expected in the case of a completely random and independent network generation process.[11]

[11] There are 10,588 directed links possible in the network data, of which 620 directed links have been formed. The ratio of these two number, 0.059, would be the expected directed clustering coefficient in case the network had been created by a process in which each node randomly and independently forms a link to any other node j with probability p=0.059.

This section has provided a detailed discussion of the Network Data, including degree distribution, density, diameters of networks, clustering of links, and persistence of reported links over time. The next two sections will discuss how these network characteristics can be used to infer more about the possible network generation mechanism which lead to the emergence of the observed line chief networks.

4.4 Network Formation: Basic Random Graph Model

In this section, I will start applying basic random network formation models to the line chief network data, to understand to what extent these models can generate the observed levels of density and clustering, and the empirical degree distribution. I will then, in the next section, turn to more complex models, which can incorporate line chief characteristics, to understand if this methodology can add to the understanding of this process. Finally, given that from Factory 1 and 2, I have two rounds of networks surveys, collected roughly one year apart, with a number of additional line chiefs having joined the network in the meantime, I can go one step further by observing the network formation process over time directly, thereby validating or rejecting some results that came out of applying standard network formation methodology.

4.4.1 The Erdos and Renyi [1959] Model

The canonical random network model was introduced by Erdos and Renyi [1959]. In its simplest static version, consider a network with n nodes, and links from node i to j are formed with probability p, independently of links between nodes i and k. The properties of random graphs are mostly studied in the limiting case if n goes to infinity. In the canonical random graph model, it is assumed that the average degree of a node stays constant as n goes to infinity, or $np = C$ for $n \to \infty$. This network formation process leads to a Poisson distribution for degrees d_i across nodes i:

$$f(d_i = k) = \frac{(np)^k e^{-np}}{k!} \tag{4.4}$$

This model is most useful as a benchmark, which demonstrates that such a simple random network fails to capture two key stylized facts observed with most empirical networks, sparsity and relatively high clustering of links (Chandrasekhar [2015]). Recall that sparse networks are defined as networks in which the average degree increases less than proportionally with n as n goes to infinity. Thus, the canonical

Table 4.8: Density and Clustering

		(1)	(2)	
Factory	Year	Density: $\hat{p}$	Clustering Coeff.	n
1	2013	0.040	0.34	57
1	2014	0.054	0.56	53
2	2013	0.250	0.44	16
2	2014	0.286	0.45	15
3	2014	0.036	0.38	60
4	2014	0.142	0.56	26
All		0.058	0.46	220

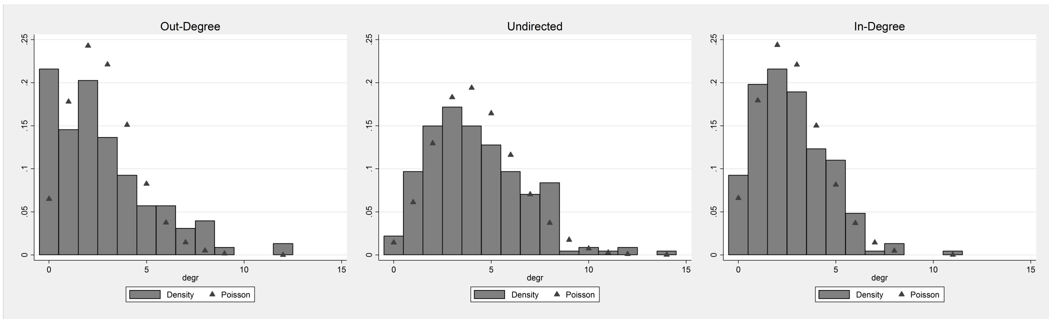
Notes: Table compares for each network on the factory-year level the empirircal network density against the empirical clustering coefficient, as defined in equation 4.3.

random graph model, in which the average degree np stays constant as n goes to infinity, creates sparse networks. However, while this model can generate the observed low density of empirical networks by choosing an appropriate low value for p, it cannot explain at the same time the high clustering of these networks. Static random graphs have a clustering coefficient of p as well. Thus, assuming constant degree, their clustering coefficient goes to zero as well as n goes to infinity.

We can use the empirical density $\hat{p}$ of networks as an estimate for p when fitting the canonical random graph model. Table 4.8 therefore compares the (directed) clustering coefficient as predicted by this model, $\hat{p}$ (column 1), against the empirical clustering coefficient, as defined by equation 4.3 (column 2). The clustering coefficient as predicted by the fitted Erdos-Renyi model is consistently below the empirical one, and by magnitudes so for all factories except Factory 2. Thus, as with many empirical networks, the static random graph model cannot generate both the observed levels of sparsity and clustering at the same time.

To test how well the predicted Poisson degree distribution matches the empirical degree distribution, Figure 4.2 plots again the degree distribution, with the fitted Poisson distribution plotted over the histograms. The left-hand graph in Figure 4.2 does so for the out-degree, the centre graph for the undirected degree, and the right-hand graph for the in-degree, using the empirical density of the network

Figure 4.2: Empirical Degree Distrubtions and Poisson Distribution



Figures show histograms of out-degree (left), undirected degree (center), and in-degree (right), and fitted Poisson distributions as predicted by the random graph model from Erdos and Renyi [1959]. Poisson distributions are fitted using the empirical density across all networks.

as estimate for p in equation 4.4. While the Poisson distribution seems to fit the undirected and in-degree distribution reasonably well, it seems to be at odd with the distribution of out-degree, which has many more nodes in the left, but also in the right tail as predicted by the Poisson distribution.[12] This is another commonly observed conflict between the predictions of the canonical random graph model and many empirical networks.

4.4.2 Growing Random Graph Models

To address the shortcoming of missing tails in the predicted degree distribution, growing random graph models have been developed. In uniformly growing random graph models, one additional node is 'born' at each time period t, and will, at the time of its birth, form m links with the already existing nodes. It can be shown that as t, and thus n, goes to infinity, the cumulative degree distribution will approach a (variation of an) exponential distribution:

$$F(d_i < k) = 1 - e^{\frac{-k}{m}+1} \tag{4.5}$$

Note that this distribution does not depend on the time that has passed in the network formation process. The expected share of nodes with degree smaller than

[12]Formal Pearson or likelihood chi-square tests reject the equivalence of all three degree distributions with the fitted Poisson distributions on at least the 2% significance level. However, the test might be too strict in rejecting equivalence of the distributions for the intention of Figure 4.2, which is to show whether the empirical distribution fits qualitatively better with a Poisson degree distribution, as predicted by the Erdos and Renyi (1959) model, or rather with exponential or power distributions, as predicted by other network models.

k will be constant as the formation process evolves. Note, furthermore, that, as each node starts with m nodes at birth, the distribution function is only defined for $k \geq m$. The mean degree in this model is $2m$ (as each node brings m links into the graph, and each link involves two nodes).

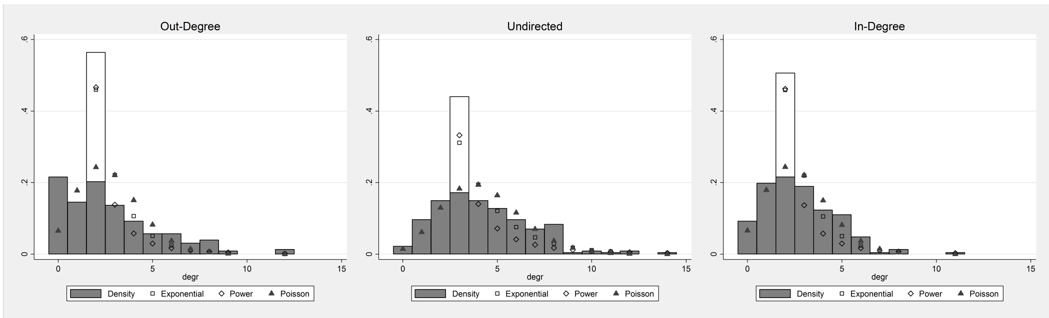
Uniformly growing random graph models have been extended to preferential attachment models (Barabasi and Albert [1999]). Now each new born node does not form m links with randomly chosen existing nodes, but to existing nodes with a probability proportional to the degree of the existing nodes. With this network formation process the expected degree distribution approaches a power distribution as n goes to infinity:

$$F(d_i < k) = 1 - m^2 k^{-2} \tag{4.6}$$

As the number of links formed is still the same as in the uniform growing random graph model, the mean degree of networks is still $2m$. Figure 4.3 plots the same degree distributions as shown already in Figure 4.2 against fitted exponential and power distributions. The parameter m was approximated by the mean degree divided by two, as predicted by these models. As both growing random graph models do not predict any node with degree less than m, the exponential and power distributions are not defined for this range in Figure 4.3. For reference, the Poisson distribution as predicted by the static random graph model is included again, the same which was already included in Figure 4.2. For the undirected and in-degree distribution, the exponential and power distribution do not seem to fit the empirical distributions better than the Poisson distribution, which as the only one of the three captures the left tail rather well. For the out-degree distribution, both the power, and especially the exponential distribution seem to capture the right tail better than the Poisson distribution. However, as they are not defined for degrees smaller than m, they cannot be tested on how well they capture the left tail. However, if we assume that the networks were indeed generated by a growing random graph process, we could assume that the observed degrees of nodes lower than m are due to underreporting of ties, and we could assume that in reality these nodes have degree m. In this case, the histograms would be bunched at degree m, as shown by the hollow bar in the graphs of Figure 4.3, and the bars to the left of the hollow bar in each graph would disappear. This hypothetical distribution would be captured better by the power or exponential distribution.[13]

[13]However, formal Pearson-chi square tests reject again equivalence of the empirical distributions with the exponential or power distribution. However, the same qualification to this rejection as

Figure 4.3: Empirical Degree Distributions and Exponential, Power, and Poisson Distribution



Figures show histograms of out-degree (left), undirected degree (center), and in-degree (right), and fitted Poisson, Exponential, and Power distributions, which are fitted using the empirical density across all networks.

To conclude, the Poisson distribution does a decent job in capturing the undirected degree distribution, and especially the in-degree distribution, and growing random graph models do not seem to add value in explaining these networks. The out-degree distribution seems to be best captured by an exponential distribution, or at least its right tail. However, in uniformly growing random graph models, the high-degree nodes are the longest existing nodes in the graph. But as shown in Table 4.4, senior line chiefs do not report more out-going links, contradicting a basic prediction of this models. Thus, growing random network models do not seem to have explanatory power for the formation of the line chief network, at least above and beyond static random graph models, which themselves do not capture the out-degree distribution well.

4.4.3 'Meeting Friends' Models

While standard growing random graph models can capture degree distributions with fat tails, they do not generate higher clustering than static random graph models

stated in footnote 12 should apply. Apart from that, the hypothetical degree distribution would obviously also have a higher mean degree, implying a higher value m (which is equal to mean degree divided by two), which would affect the fitted exponential and power distributions again. As I do not want to overemphasize the aspect of these bunched degree distribution, I left the non-adjusted power- and exponential distribution in graph. The pictures do not qualitatively change with the adjusted distributions, as the two distributions are simply shifted slightly upwards. The formal rejection of equivalence of distributions from the Pearson chi square test would still hold. Graphs with these distributions adjusted for higher m are available from the author.

do. For n going to infinity, and with m constant over time, the clustering coefficient goes to zero, just as for static random graph models. As n grows, the likelihood that when matching randomly with m out of the n nodes, some of the m nodes are connected among themselves goes to zero. It can be shown that this even holds in case of preferential clustering, even though it is difficult (Jackson [2008]). To capture non-negligible clustering in a random graph framework, it has to be enriched with more structure, as done in the 'meeting friends' model of Jackson and Rogers [2007]. In this growing random network model, each new-born node forms again m links to existing nodes. However, while m_r of these links are formed through random matching with existing links, the $m - m_r$ remaining links are formed to randomly chosen neighbours of the m_r nodes connected to in the first step. It turns out that in this model it is easier to derive predictions for the degree distribution when considering only directed links. In this case, the probability of a node i being matched to a new born node only depends on the age of node i, not on the one of its neighbours, as each node has the same number of outgoing links that could go to node i. For directed networks, the predicted in-degree distribution is then (out-degree being m for each node):

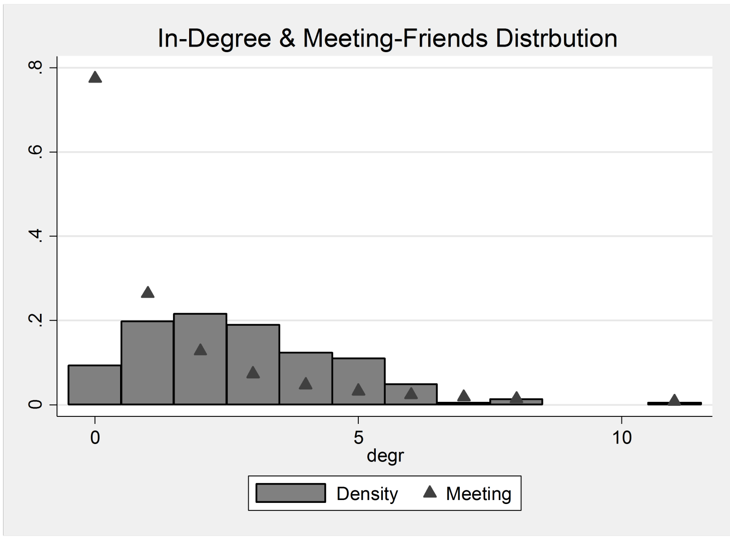
$$F(d^{in} < k) = 1 - \left(\frac{mm_r}{m - m_r}\right)^{\frac{m_r}{m-m_r}} \left(d^{in} + \frac{mm_r}{m - m_r}\right)^{\frac{-m_r}{m-m_r}} \tag{4.7}$$

This is an augmented power distribution. Given that, as already shown, the in-degree distribution in the line chief network is well approximated by a Poisson distribution, it will not be approximated well by the distribution derived from the meeting-friends network. Figure 4.4 plots this distribution over the empirical in-degree distribution to show the mismatch. Therefore, it does not appear that the 'meeting friends' class of random graph models can be used to explain first the observed degree-distribution, and subsequently the high correlation coefficient of the network.

4.4.4 Block Random Graph Models

The attempts so far to capture both the network density and clustering levels of the line chief network with random graph models ignored the fact that ties are reported disproportionally to line chiefs from the same floor, a tendency which automatically increases clustering. We can incorporate floors in random graph models using block models, which date back to Holland et al. (1983). In these models, all nodes in a network are allocated to blocks B, a mutually exclusive and exhaustive partition of

Figure 4.4: Degree Distributions in 'Meeting Friends' Model



Figures show histograms of in-degree, and the Power distributions from the meeting friends model of Jackson and Rogers [2007] as shown in equation 4.7, with m fitted with the empirical density of the networks, and $m_r = 0.95m$.

the nodes. The probability that a link is formed between nodes i and j then depends on the pair-block B_{rs} which contains i and j, with pair-block B_{rs} being the pair of blocks B_r and B_s, such that B_r contains node i and B_s node j. Essentially, nodes that are located in the same block have a distinct probability that ties are formed between them (which itself can vary from block to block), compared to nodes that are not in the same block (where the probability can also vary across the pairing of two blocks that contains the nodes).

The concept of the blocks naturally lends itself to sewing floors in the line chief network. I estimate a restricted version of the block model with the data from the network, in which I only consider two different probabilities; probability p^S for a link being formed between two line chiefs from the same floor (regardless of which floor), and probability p^R for a link being formed between two line chiefs from different floors (regardless of the pair of floors). I estimate this restricted version to see whether simply the within/across floor variation in network densities can account for the observed levels of clustering. The probabilities of link formation in the block model between two nodes in block r and s can be estimated by the pair-block density $\hat{p}_{rs}$, the density of the network in which each (potential) link connects a node from block r with a node from block s, whereby $r = s$ or $r \neq s$. Table 4.9 shows the estimated probabilities $\hat{p}^S$ and $\hat{p}^R$ for each factory-year network. Between 23% and 68% of all possible directed links within floors are reported in the networks, with the average of 36% across the factories (column 1 – average weighted by number of possible within-floor links per factory-year). For Factories 2 and 3, this estimate for p^S are now very close to the within-floor clustering coefficient, as shown in column 3, while for Factory 1 it is half of the coefficient and roughly two-thirds for Factory 4.[14] This is a marked improvement compared to Table 4.8, where the overall clustering coefficient was ca. eight times as large as the overall density. Cross floor density varies between essentially zero at Factory 2 in 2014 (only two links reported across the six floors with 53 line chiefs) to 13% at Factory 2. This density is well in line with cross-floor clustering at Factory 1, 2 (at least in 2013) and 4, while out of line in Factory 3. However, the discrepancies between cross-floor density and clustering at Factory 3 could be due to small sample bias.[15] Therefore, it seems that static block random graph models can capture both density and clustering of the line chief networks well.

To conclude this section on basic random graph models, as in many empirical settings, the basic random graph model due to Erds and Renyi (1959), and subsequent growing random graph models struggle to model both the observed density and clustering at the same time. Also variations of the model explicitly designed to reconcile sparsity and high clustering, such as the 'meeting-friends' model from Jackson and Rogers (2007), do not provide a good fit, as their implied degree distribution is at odds with the empirical one. However, a simple block random graph model gets close to reconciling observed network density and clustering.

This result contributes to a debate on the reasons for the ubiquitous high levels of clustering observed in empirical networks. Is it due to homophily, meaning that nodes of similar characteristics have a higher likelihood of forming links, thus generating clusters among themselves? Or is it due to network externalities, such that links which are part of a cluster yield higher utility? For example, friendship with another node could be more enjoyable if one shares third friends with this

[14]Within floor clustering coefficient refers to the directed clustering coefficient, as defined in equation 4.3, on the network that ignores any pairs of nodes which are not on the same floor.

[15]At Factory 3, only 17 out of 3,314 possible cross floor links were reported, and there are only three instance in which one directed cross-floor link can be followed by another cross-floor link from that node to a third node on yet another floor. In one of these three cases, however, (see graph of Factory 3 in the Appendix E), the first node does also have a direct node to the third node, resulting in a directed clustering coefficient of 0.33.

Table 4.9: Block-Random Graph Models

		(1)	(2)	(3)	(4)
		$\hat{p}^S$	$\hat{p}^R$	Clustering	Clustering
Factory	Year	(within flrs.)	(across flrs.)	within flrs.	across flrs.
1	2013	0.231	0.004	0.414	0.000
1	2014	0.334	0.001	0.577	0.000
2	2013	0.617	0.128	0.590	0.093
2	2014	0.683	0.127	0.622	0.226
3	2014	0.496	0.005	0.461	0.330
4	2014	0.412	0.018	0.649	0.000
All		0.359	0.009	0.516	0.126

Notes: Table compares for each network on the factory-year level the empirical (directed) network density within floors ($\hat{p}^S$) and across floors ($\hat{p}^R$) against the empirical (directed) clustering coefficient as defined in equation 4.3, calculated within and across floors

node. These two scenarios are empirically often difficult to discriminate (Graham [2015a]). However, in the line chief network, in which we observe very high levels of clustering, we do have a simple measure of homophily, line chiefs working on the same floor. Once we allow for the possibility of differential likelihood of social ties being formed within floors, static random graph models, which ignore network externalities, explain the observed level of clustering. It could be that also in other empirical networks with high clustering, if we had better information on underlying group structures of nodes, a lot of variation in clustering could be captured by block random graph models.

4.5 Network Formation: Heterogeneous nodes models

In the basic network model, all nodes are homogeneous in the sense that links form with equal probability between two nodes. Growing random graph models already introduced heterogeneity in terms of 'age' and degree of nodes which affect the likelihood of them forming links with other nodes, while block random graph models assign nodes to certain groups with the probability of links forming between two nodes depending on each's group. However, more recent models allow for much richer heterogeneity on the level of individual nodes i, by incorporating observable node characteristic vectors x_i.

One early generation of these models came in form of dyad-level regressions, in which a dummy for whether a link exists between two nodes was regressed on absolute and relative values of the elements in the vectors x_i of the pair of nodes (Fafchamps and Gubert [2007a,b]; Comola [2012]; Graham [2015b]). For the sake of estimateability, this literature ignores externalities of link formation, which is at the heart of many strategic network models. It assumes that the utility a nodes i derives from a link with node j is separable from the existence of any other links node i or j have. This still leaves open the question of correlation of errors within and across nodes. As unobserved characteristics of nodes that form links are likely correlated, so too would be their errors. One way this literature addresses this problem is by clustering standard errors on social distance of nodes (Fafchamps and Gubert [2007a,b]).

A recent extension of the early dyad level-models are Sub-Graph Models (SUGMs, Chandrasekhar and Jackson [2014]). These models assume that nodes cannot only agree to form bilateral links, but any kind of sub-graphs, such as triangles or cliques of any other number of nodes which are fully connected among each other. In an expositional version of the model, Chandrasekhar (2015) assumes that in a first step, all possible triplets of nodes i, j, and k meet and decide whether to form a triangle among them (that is, three links ij, jk, and ik), depending on the node characteristics, both in absolute terms as in relative terms to each other. In a second step, all not-yet connected pairs of two links meet, and decide whether to form bilateral links (also referred to as 'unsupported' links). The explicit introduction of triangles or higher-order sub-graphs allows the model to capture arbitrary levels of clustering in otherwise sparse networks. The estimation of the model is straightforward in a regression framework if the network is sufficiently sparse, as in this case the number of 'incidental' triangles, which were not formed explicitly as triangles but by a combination of otherwise formed triangles and bilateral links tends to zero as the number of nodes goes to infinity. Thus, they do not bias the estimates of the parameters which govern triangle and unsupported link formation. Chandrasekhar and Jackson [2014] show that the sparsity condition holds if every node in the network participates in at most $o(\sqrt{n})$ unsupported links and $o(\sqrt{n})$ triangles, which is fulfilled by most empirical networks.

Table 4.10 below shows the results when estimating simple dyad-level and

Sub-Graph Models from the line chief networks. Columns 1 and 2 show the results from the dyad-level model, where the sample consists of all directed pairs of nodes i and j in the networks, and the dependent variable is a dummy whether a directed link runs from node i to node j. The econometric model estimated in column 1 and 2 of Table 4.10, estimated using logit regression, is:

$$link_{ijf} = \alpha + \beta^S(x_{if} + x_{jf}) + \beta^D(x_{if} - x_{jf}) + \beta^l\gamma_{ijf} + \delta_f + \epsilon_{ijf} \tag{4.8}$$

In this model $link_{ijf}$ stands for a link from node i to node j in factory f, x_i is a vector of observable characteristics of node i, γ_{ijf} represents relationship characteristics of the two variables, such as whether they are on the same floor, and δ_f are factory fixed effects, while ϵ_{ijf} is an independent random term. Covariates of node i and j need to enter dyadic regressions in a symmetric way, which, with directed network data, is achieved best by assuming one coefficient for the sum of x_i and x_j, and one for the difference between them (Fafchamps and Gubert [2007a]). In principle, node level fixed effects δ_{if} can be added as well, as each node is part of $2(n-1)$ directed dyadic relationships. However, Graham [2015b] shows that in this case, the variance of the estimator is inflated relative to standard asymptotic confidence intervals, and needs to be adjusted. Therefore, in this demonstration of the model, node fixed effects are not included. I cluster standard errors on social distance, with a distance cut-off of one link. This means that I allow arbitrary correlation of the residual from the observation of link i to j with the residuals of all dyad-level observations which include either node i or j.[16]

Column 1 from Table 4.10 shows the results when estimating equation 4.8 using the data from all six networks on the factory-year level, and using all node pairs within a factory-year as observations. The results do not show any effect of the four included line chief characteristics age, seniority as line chief, seniority in factory, and education. Only a dummy for working in the same floor has a strongly positive effect, in line with prior results.

The formation of links within floors might have different determinants than those across floors. Given that the large majority of links are formed within floors, it might be instructive to only look at the determinants of links within floors, which is done in column 2, Table 4.10. As all node pairs in the sample are now located on the same floor, the same-floor dummy is dropped. In this specification, the sums

[16] I implement this clustering approach in STATA using the 'ngreg' command from Marcel Fafchamps homepage: http://web.stanford.edu/ fafchamp/resources.html

of seniority as line chiefs has a positive effect on link formation, while the difference has a negative one. This implies that, within floors, links are disproportionally formed from more junior to more senior line chiefs in terms of years the line chief already works on this position in the factory, and is in line with the results from Table 4.4 which show that these senior line chiefs has a higher in-degree but not higher out-degree.

Columns 3-6 show the results of estimating the sub-graph model with triangles first and then remaining node pairs. While dyad-level regression models can be readily estimated using either directed or undirected network data, it is easier to estimate this sub-graph models using undirected network data, as it is more straight-forward to determine whether an undirected link is part of an undirected triangle. Thus, columns 3-6 work with the undirected network again. Following Chandrasekhar [2015], column 3 shows the results from the first step of estimating the sub-graph model, a logit regression using the sample of all 90,521 possible undirected triangles between node i, j, and k within the six networks, of a dummy tr_{ijkf} for a triangle being formed on the characteristics of line chief i, j, and k:

$$tr_{ijkf} = \alpha + \beta^S(x_{if} + x_{jf} + x_{kf}) + \beta^D(|x_{if} - x_{jf}| + |x_{if} - x_{kf}| + |x_{jf} - x_{kf}|) + \beta^T\gamma_{ijkf} + \delta_f + \epsilon_{ijf} \quad (4.9)$$

While β^S is again a vector of coefficients for the sum of the covariates, now across all three nodes involved in a link, β^D is a measure for the role of homophily in link formation, the often observed tendency of nodes to form links with nodes of similar characteristics. A negative sign for β^D would indicate homophily. Using the absolute value of the differences of node characteristics ensures that the node characteristics are included in the necessary symmetric way when using undirected network data. The term γ_{ijkf} now captures triangle level characteristics, such as whether all three nodes are on the same floor, while, again, factory fixed effects are included. However, looking at the results from column 3, there is no evidence for homophily in this network. Again, there is a tendency for more senior line chiefs to form triangles, but, if anything, diversity in seniority increases the likelihood a triangle is filled. Furthermore, as usual, nodes from the same floor have a strongly increased likelihood to form triangles. Column 4 repeats the regression from column 3, while restricting again the sample to all possible triangles within the same floor. While the pattern for seniority as line chiefs holds in this sample as well, a similar

Table 4.10: Dyadic Regression Model and Sub-Graph Model

	(1)	(2)	(3)	(4)	(5)	(6)
MODEL	Dyad-Regression		Sub-Graph-Model			
VARIABLES	Link	Link	Triangle	Triangle	Rem. Link	Rem. Link
Diff. Age	-0.018	-0.013				
	(0.011)	(0.012)				
Diff. Sen.ity as LC	-0.049	-0.076**				
	(0.040)	(0.038)				
Diff. Sen.ity in Fact.	-0.039	-0.028				
	(0.028)	(0.023)				
Diff. Education	-0.008	0.008				
	(0.026)	(0.029)				
Sum Age	-0.020	-0.016	-0.015	-0.015	0.000	0.049
	(0.016)	(0.018)	(0.016)	(0.019)	(0.030)	(0.044)
Sum Sen.ity as LC	0.069	0.106**	0.070*	0.077*	-0.028	-0.048
	(0.046)	(0.049)	(0.041)	(0.045)	(0.069)	(0.100)
Sum Sen.ity in Fact.	0.040	0.004	0.008	-0.024	0.033	-0.048
	(0.034)	(0.033)	(0.027)	(0.028)	(0.043)	(0.052)
Sum Education	0.018	0.022	0.035	0.066*	0.005	0.000
	(0.029)	(0.030)	(0.038)	(0.038)	(0.085)	(0.132)
Hom. Age			-0.005	0.007	0.037	0.062
			(0.013)	(0.014)	(0.043)	(0.065)
Hom. Sen.ity as LC			0.056*	0.068**	0.146*	0.294*
				(0.031)	(0.084)	(0.153)
Hom. Sen.ity in Fact.			-0.009	0.021	-0.084*	-0.032
			(0.021)	(0.020)	(0.047)	(0.093)
Hom. Education			0.025	0.053*	-0.079	0.034
				(0.031)	(0.072)	(0.090)
Same Floor	4.496***		7.122***		3.00***	
	(0.239)			(0.426)	(0.380)	
Constant	-5.704***	-1.466	-10.83***	-5.488**	-5.610*	-5.633
	(1.219)	(1.267)	(2.345)	(2.296)	(3.117)	(4.347)
Observations	9,540	1,366	76,930	1,547	4,265	332
Factory FE	YES	YES	YES	YES	YES	YES

Notes: Column 1 and 2 show results of regressing a dummy for the existence of a link on the node-pair level on the sums and the differences of observable characteristics of the pair of line chiefs. Column 2 does so only for pairs of line chiefs on the same floor. Columns 3 and 4 regress the existence of a triangle of links between three nodes on the sum, and the sum of the three absolute differences of the characteristics in each of the three pairs involved in the triplet (see equation 4.9). 'Sen.ity as LC' measures months line chief works already as line chief in factory, and 'Sen.ity in Fact' month line chief works on any position in factory. Education is IPA Bangladesh specific ordinal measure of educational attainment. Robust Standard Errors in Column 1 and 2 clustered on social distance, distance cut-off one link: *** p<0.01, ** p<0.05, * p<0.1

pattern now also emerges for education. More educated line chiefs have a higher likelihood to be part of triangles, which however involves also taste for diversity in education when forming triangles.[17]

Column 5-6 represent the second step in the estimation of this simple triangle-dyad subgraph model. Using the sample of all remaining node-pairs which are not part of an actual triangle in the network, a logit regression of the following empirical model is run:

$$link_{ijf} = \alpha + \beta^S(x_{if} + x_{jf}) + \beta^D(|x_{if} - x_{jf}|)| + \gamma_{ijf} + \delta_f + \epsilon_{ijf} \tag{4.10}$$

This model resembles the dyad-regression model shown in equation 4.8, but adapted for undirected network data, and estimated now using only the sample of pair-nodes which are not part of a triangle. The coefficient vector β^D now again measures the effect of homophily on bilateral link formation. The regressions reveal again a slight 'heterophily' in terms of seniority both when estimated on the overall sample of remaining node pairs (column 5) or only using within floor pairs (column 6). However, in the overall sample, also homophily with respect to seniority in the factory becomes visible.

4.6 Actual Network Formation

To conclude the discussion on network formation, I can make use of the fact that I have two rounds of survey data available from Factory 1 and 2, conducted roughly one year apart from each other. This allows me to study how those line chiefs that newly joined the ranks of line chiefs between the two survey rounds formed ties with both the existing as well as the other newly joined line chiefs. The outcome of this analysis can then be compared against the assumptions of network formation models, or their results when estimating these models, which attempt to infer the past network formation process from 'snap-shots' of network data at a given point in time.

In principle, as the network data includes for each line chief the months he or she already worked in the factory and as line chief, one could simply study which ties those line chiefs report that joined the ranks of the line chiefs recently. To some extent, this is already done by the analysis shown in Table 4.10, which showed

[17] Education is strongly negatively correlated with seniority in factory, somewhat positively correlated with seniority as line chief, and uncorrelated with age.

that less senior line chiefs who have joined recently have a tendency to report links to line chiefs that already worked longer on this position. The availability of two rounds of network data, however, allows to specifically extend the analysis to study whether newly joined line chiefs preferentially form ties with existing line chiefs that already have many ties. This is, for example, the central assumption from preferential attachment random graph models, as discussed above. However, when regressing a dummy for a link between two line chiefs on the degrees of the involved line chiefs, the degrees would be mechanically correlated with the dependent variable. We could not disentangle whether the link was formed to the line chief because he had a high degree, or because of some innate characteristics that made the line chief more popular to form links with at that point in time. Therefore, I instead use lagged in-degree of the line chief from the previous survey from 2013 as independent variable, to test whether newly joined line chiefs by 2014 preferentially form links to line chiefs who had a high degree in 2013. Section 3 has shown persistence of ties reported across the two survey rounds, and the correlation coefficient of in-degree in 2013 and 2014 is 0.52.

Table 4.11 shows the results when studying the determinants of links formed by the 'new' line chiefs. It thus only uses data from the second survey from 2014 from Factory 1 and 2, in which 26 line chiefs were interviewed that were not interviewed the year before. Table 4.11 focuses on the links that these 26 new line chiefs reported, both with line chiefs that were already interviewed the previous year, and with the other new line chiefs. In column 1, Table 4.11, I estimate a dyad-level regression model as shown in equation 4.8, using the sample of directed pairs of line chiefs i to j, in which i is a new line chief, and j can be any other line chief. Column 1 is the baseline model, regressing a dummy for reported links by new line chiefs to other line chiefs on a dummy for whether the other line chief was already a line chief last year, whether he or she works on the same sewing floor, and on the sum and differences in the four main line chief observable characteristics. While, as usual, working on the same unit has a strongly positive effect on link formation, there is no significant effect of forming links with line chiefs that were already present in the previous survey round. Instead there seems to be a strong tendency to form links to older line chiefs. This is somewhat at odd with the results from the previous table which showed a preference to form links to more senior line chiefs, but not older line chiefs. While this difference could be driven by the restriction of this sample to Factory 1 and 2, it could also be that, while new line chiefs have a desire to form ties to more senior line chiefs, they cannot readily find out the seniority, or social ‘status’, of all line chiefs in the factory. Instead, they use age as a more readily observable proxy for status to form initial ties.

Column 2 regresses link formation on the central variables of interest, in- and out-degree of line chiefs from the previous survey in 2013 at the factories. It thus restricts the sample to pairs of line chiefs in which one line chiefs has not been a line chief in the factory in the previous survey while the other has been. There seems to be indeed a tendency to form ties to line chiefs that had a higher in-degree in 2013. This effect is not explained by the inclusion of the line chief observables again in column 3, which, if anything, slightly strengthens the result (p-value of in-degree: 0.050). The central assumption of preferential attachment network models seems to be confirmed in this data, even though the model’s predicted degree distribution did not fit with the empirical one from the line chief network. Out-degree in 2013, on the other hand, has no effect on link formation. The effect of differential age on link formation remains unchanged compared to column 1.

Finally columns 4 and 5 repeat column 2 and 3, but now regressing not a dummy for a new line chief reporting a link to an existing one, but instead an existing one reporting a link to a new line chief, on the observable characteristics of both line chiefs. Neither in- nor out-degree of the existing line chief in 2013 has an effect on whether a tie is reported to a new line chief. What does show an effect is seniority as line chief; less senior existing line chiefs are more likely to report link to new line chiefs than more senior ones.

4.7 Degree as Predictor for Turnover

To conclude this chapter, the fact that two rounds of surveys are available from Factory 1 and 2 will also be used to study whether the number of ties a line chief has in the first round predicts whether the line chief will still be present at the factory in the survey one year later. Relating turnover behaviour of employees to network centrality measures has been done by a number of studies in organizational research (Feeley and Barnett [1997]; Feeley et al. [2008]; Mossholder et al. [2005]). Interestingly, these studies have produced to some extent conflicting results on which measure of centrality matters for predicting turnover of workers. While in Mossholder et al. [2005]’s dataset in-degree of a worker is significantly negatively correlated with subsequently leaving the organization, in Feeley et al. [2008] out-degree is, while

Table 4.11: Actual Network Formation

VARIABLES	(1) Out-Link	(2) Out-Link	(3) Out-Link	(4) In-Link	(5) In-Link
Existing LC	0.161				
	(0.493)				
Same Floor	4.761***	4.194***	5.153***		
	(0.656)	(0.648)	(0.749)		
In-Degree '13		0.179*	0.197*	-0.132	-0.128
		(0.097)	(0.101)	(0.113)	(0.202)
Out-Degree '13		0.036	0.026	-0.013	-0.043
		(0.068)	(0.081)	(0.100)	(0.129)
SUM Age	-0.084***		-0.125***		-0.031
	(0.030)		(0.046)		(0.040)
SUM Seniority as LC	0.236		0.258		-0.247*
	(0.256)		(0.249)		(0.132)
SUM Seniority in Fact.	-0.050		-0.087		0.058
	(0.126)		(0.141)		(0.068)
SUM Education	-0.137		0.010		-0.025
	(0.134)		(0.137)		(0.098)
DIFF. Age	-0.088***		-0.125***		0.011
	(0.030)		(0.040)		(0.041)
DIFF. Senior. as LC	0.085		0.141		0.291**
	(0.217)		(0.238)		(0.146)
DIFF. Senior. in Fact.	-0.173		-0.137		0.042
	(0.130)		(0.145)		(0.076)
DIFF. Education	-0.186		-0.180		-0.091
	(0.121)		(0.130)		(0.128)
Constant	2.986	-5.310***	0.446	-0.964**	2.856
	(4.839)	(0.631)	(5.667)	(0.418)	(4.494)
Observations	904	686	608	117	105
Factory-YEar FE	YES	YES	YES	YES	YES

Notes: Column 1 regresses a dummy on whether a line chief, who was not present in the 2013 survey, reports a social tie to another line chief on the sum and difference of his or her characteristic and those of the other. Column 2 and 3 include lagged degree from the 2013 survey from the other line chiefs, thereby restricting the sample on possible links reported by new line chiefs to those which were already present in the survey from the previous year. Column 4 and 5 regress a dummy for a tie reported by a line chief who was already present in the 2013 survey to a line chief who was not on sum and differences of the line chief characteristics. Regressions control for factory-year fixed effect. Robust standard errors clustered at level of line chief reporting tie in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

in-degree is not (Feeley and Barnett [1997] use degree in an undirected network). Therefore, the line chief data from Bangladesh could add valuable additional evidence from workers with a different cultural and socio-economic background. While in-degree could capture the status of a member of a network, and thus the resources he or she could muster to deal with difficulties on the job, out-degree could capture the engagement of a person with his organization.

Table 4.12 below shows the results when regressing, among the sample of line chiefs interviewed in Factories 1 and 2 in 2013, a dummy of whether the line chief is still present in the 2014 survey on his or her out- and in-degree in 2013, and on several other observable line chief characteristics. In line with Mossholder et al. [2005], but not with Feeley et al. [2008], I find that in-degree is negatively correlated with turnover, but not out-degree, after controlling for all available line chief observables and factory fixed effects. This effect becomes especially strong when using weighted in-degree, in which an incoming link is weighted by the number of different dimensions it is reported on (spending lunch break together, visited home, friendship,...). Also undirected degree is not significantly related with turnover. This speaks in favour of the hypothesis that it is specifically the 'social resources' one has in the factory that matter more for the decision to stay or leave than one's social engagement with the organization.

4.8 Conclusion

This chapter has introduced a novel social network dataset among line supervisors from four Bangladeshi garment factories. It has tested and estimated several network formation models, to understand which class of these models could provide most insight into the formation process of these networks. Two main results remain from this analysis. First, static block-random network models do a good job in capturing both the observed levels of density and clustering of the network data, something that both the canonical random graph models from Erdos and Renyi [1959], as well as growing random graph models, often struggle to achieve. Second, as common in work-place social networks, the network exhibits a core-periphery pattern, with more senior line chiefs forming the core, and more junior line chief having a preference to form ties to the line chiefs in the core as compared to other line chiefs.

Especially the availability of two rounds of network surveys from two of the four factories in the study allows the derivation of some additional results. First, there is significant persistence of ties being reported across one year. 51% of ties being reported in 2013 are reported again in 2014. Had ties in 2014 been formed randomly and independently from ties in 2013, only 7.5% of ties from 2013 should have been reported again by chance in 2014. There is very little evidence so far in the literature on persistence of reported social connections. Second, we could see that in-degree of line chiefs in 2013 had predictive power on whether newly joined line chiefs in the 2014 survey report social ties to those line chiefs already present in the survey from 2013. This confirms a central assumption from preferential attachment random network models. Finally, in-degree, but not out-degree of a node is positively correlated with still working at the factory one year later, which is additional evidence for the empirical literature on turnover in organizational studies, which so far has found conflicting evidence on which kind of network position measure matters for predicting worker turnover.

Table 4.12: Degree and Turnover Likelihood

VARIABLES	(1) Stay	(2) Weighted Deg. Stay	(3) Stay	(4) Weighted Deg. Stay
Out-Degree	0.020	-0.024		
	(0.072)	(0.030)		
In-Degree	0.268*	0.124**		
	(0.159)	(0.061)		
Undir. Degree			0.130	0.014
			(0.087)	(0.023)
Seniority as LC	-0.213	-0.198	-0.148	-0.106
	(0.178)	(0.165)	(0.146)	(0.138)
Seniority in Fact.	0.116	0.123	0.114	0.108
	(0.118)	(0.115)	(0.104)	(0.100)
Age	0.051	0.046	0.061	0.055
	(0.047)	(0.048)	(0.046)	(0.046)
Education	0.111*	0.109*	0.103	0.107*
	(0.065)	(0.064)	(0.065)	(0.063)
Started as LC	0.024	0.015	0.159	0.221
	(0.551)	(0.545)	(0.515)	(0.505)
Constant	-3.511*	-3.156*	-3.872**	-3.556*
	(1.855)	(1.871)	(1.871)	(1.829)
Observations	67	67	67	67
Factory FE	YES	YES	YES	YES

Notes: Table shows results when regressing among all line chiefs surveyed in 2013 in Factory 1 and 2 a dummy on whether he or she is still present at the factory one year later on his or her out-degree and in-degree (columns 1 & 2), or undirected degree (column 3 & 4) in 2013, and all other observables. 'Seniority as LC' measures month line chief worked as line chief in factory, and 'Seniority in Fact' months working on any position in factory. Education is IPA Bangladesh specific ordinal variable of educational attainment. Regressions control for factory FE. Robust standard errors reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Appendix A

Description: Operator Training Project, Phase 1

A.1 Project Design

The first phase of the project began in November 2011. The training program was designed with the goal of increasing the number of female supervisors in factories, and GIZ expressed a preference that we train only female operators as part of the project. Recognizing the value of having some comparison sample of male operators, we agreed with GIZ to train four females and one male from each of the participating factories. We began contacting potential factories, with a letter of introduction from a large UK-based buyer, in August 2011. The first training session began in November 2011. After six rounds of training, we stepped back in January 2013 to assess the design.

Our aim was to select a sample of factories capable of selling directly to large international buyers. We obtained an initial sample using data from transaction-level trade data obtained from the Bangladeshi National Bureau of Revenue. These data provide volume (net weight) and value of exports at the shipment level. The data have identifiers which allow data from individual exporters to be aggregated. We aggregated data by exporter and calculated the unit value (U.S.$ per kilogram) for each exporter / product / year. We also summed total exports by exporter. Using these two measures, we selected a sample of firms with annual shipment volumes large enough to sell directly to large foreign buyers, with unit values in the range of mid-level buyers. This selection process yielded an initial sample of 665 exporters. We then selected the group of around 20 suppliers to one particular mid-range buyer based in the UK. For each of the 665 exporters on the initial list, we created a score based on export volume and unit values indicating how close the exporter was to the 20 suppliers of the initial UK-based buyer. We selected around 400 exporters, and searched local directories and the internet for contact information. This yielded a sample of 230 factories, which we began contacting in August 2011.

By November 2011, after contacting about 200 of the factories on the list, we had received an initial commitment to participate in the project from 96 units of 85 distinct factories. Table A1.1 shows how the characteristics of the 85 factories differ from the initial list. The table shows both a comparison characteristic by characteristic, and the p-values from a probit regression including several of the characteristics. We find that that those factories agreeing to participate sell to more buyers, and sell to higher-end buyers. The quality of buyers is measured by the average unit price paid by each buyer. For each seller, we then ordered the buyers by unit price, and measured the unit value paid by the buyer at the 90th percentile in the ranking. We also find some evidence that the participating factories had higher rates of recent growth and export products to a larger number of countries.

Participating factories were randomly placed into one of eight treatment rounds of 12 factories each. In practice we allowed factories to defer participation to a later round once, and in the end, several factories decided not to participate. By December 2012, when training round 6 began, we had exhausted the initial list of 96 factories. Note that all of the comparisons we will make with trainees control for factory fixed effects, so we view the factory-level attrition issue as mainly one of external, but not internal, validity. During the second round of the program, discussions with the local office of the International Finance Corporation led to inclusion of seven factories located in the Dhaka EPZ in the project. These factories were added in training rounds 4 and 5.

Table A.1 shows characteristics of the factories participating in rounds 1-6. The factories are large, averaging 19 production lines and 2,100 workers. Somewhat more than half of the employees in a typical factory work in the sewing section. The distributions are slightly right-skewed, with the median factory having 15 production lines, with 2,000 workers in total, of which 59% are in the sewing section. A typical factory had been operating for 12 years. Given the rapid growth of the sector, this is very likely older than the industry average.

Table A.1: Take-Up of the Program

	Signed-Up N = 85	Not Signed-Up N = 145	p-value	p-value Probit
Size (Export, 1000 Kgs)	830.4	683.8	0.11	0.44
Average Unit Value (per Kg)	925.9	883.8	0.15	0.01
Growth (Sales 2010 to 2009)*	1.89	1.46	0.08	–
Number of Destinations	10.1	8.3	0.09	0.18
Number of Buyers	9.75	8.3	0.06	0.02
Number of Products	3.01	2.91	0.32	0.31
Main Product in Woven	0.59	0.54	0.26	–
Year of first export	2006	2006.2	0.2	–
Median Buyer	560	631	0.18	0.44
90th Pcntile Buyer	183.6	283	0.03	0.01

Notes: * On a sample of 80 and 135 exporters respectively.

A.2 Selection of Trainees

Our aim was to select from each factory four female and one male operator for training, and a valid comparison group against which to measure the trainees. The details of selecting workers evolved a bit across training rounds, as we describe below, but in all rounds the process started with factories selecting a pool of potential trainees to which we administered a diagnostic test. The test was based on one designed by GIZ to measure literacy, a requirement for the training, and technical knowledge. We also gave the potential trainees a short non-verbal reasoning test and asked them questions about aspirations to work as a line supervisor. Because women were sometimes forbidden to participate in the training by their families, we also asked the potential trainees if their families would allow and support them to attend the training. Potential trainees were excluded if they did not pass the literacy test or said their families would not allow them to participate in the training.

For training rounds 1 to 3, we asked the factories to identify 16 female and 4 male operators who were good candidates for the training. We ranked the nominees according to their diagnostic score and then selected the two females with top marks on the diagnostic test as trainees. We then assigned a random number to the female trainees ranked 3rd to 6th on the diagnostic test, and assigned the two with the highest random numbers to training, and the two with the lowest random numbers to control. Among the males, we followed a similar procedure by taking the males with the top two marks and randomly assigning one to treatment and one to control.

In round 4, we modified the selection process to allow the factory to choose two females they wanted to send to training, conditional only on them demonstrating a basic level of literacy. We then took the top four females after excluding the two selected by the factory and randomly selected two for treatment and two for control. We also altered the method of replacing trainees when the selected individuals declined to participate.

In round 5, we modified the process further by reducing the number of operators the factory identified as candidates to eight females and four males. The factory then selected two of the eight females for training; the remaining two females and the male were selected randomly in the same manner as the previous rounds. We further modified the method for selecting "replacement" trainees, as described below.

There was a non-trivial amount of noncompliance. Over the six rounds, 50 workers assigned to training did not attend at all, and an additional eight attended for less than one full week. Factories most often reported that these workers either had decided they did not want to attend, or their families had said they could not attend. However, the family was most likely to intervene in the case of female trainees, while we note that the percentage of non-complying males assigned to training (21.2 percent) was higher than the percentage of non-complying females assigned to training (15.2 percent).[1] These non-compliers were replaced by 40 workers receiving training even though they were not assigned to training including 19 workers assigned as controls. Thus, non-compliance is a concern in the Phase I data when we compare the outcomes of those assigned to treatment against the controls.

As with the selection of trainees, the protocol for selecting replacements also evolved over the training rounds. In training round 1, the factories chose the replacements themselves, as we had not anticipated the severity of this non-compliance. Beginning in round 2, we insisted that the factory send the next female or male on the diagnostic ranking if a selected trainee declined to attend. Then, beginning in round 5, we altered the initial selection process to add a third female control – selecting 2 of the females ranked 3 to 7 by diagnostic score – and a second male

[1]We interpret this as suggesting that factories cared more about which males received training than they did about which females received training, perhaps because they did not plan to promote all of the females.

control – selecting one of the males in the top three diagnostic scores as the trainee. Replacements were then selected at random from among the controls.

Over the first six training rounds, 271 operators (213 females and 58 males) received training. We exclude from this total eight workers who attended for five days or fewer. Conditional on attending at least one week, attendance was very high. Out of the 36 training days, males attended 34.4 days on average and females 34.5 days. All but two of the men attended at least four of the six training weeks, as did 96 percent of the women.

After the sixth training round, we decided to suspend the training temporarily. Having already gathered a substantial amount of data and information, we felt we would gain by analysing those data and perhaps adjusting the design for the remaining factories. We resumed the training with the start of Phase 2 of the project in February 2014, which's details are described in the main body of the paper.

Appendix B

Production Data: Description and Collection

As part of both Phase 1 and 2 of the project, we collected daily production data from all factories on the sewing line level. The data is similar in its format and organization across the two project rounds. However, in Phase 1 of the project we collected data in a two week interval every other month, while in Phase 2 we collected data for each day between January 2014 and March 2015. Given the continuity and greater amount of data, we base the analysis in the main part of the paper on the data from Phase 2, which we describe in more detail in this appendix.

We collected the data with three main outcome variables in mind: line-level productivity, the quality defect rate, and worker absenteeism. We asked factories to share all internal data needed to construct these variables. The standard measure of productivity in the Bangladeshi garment industry is (piecewise output * SMV)/(workers * daily hours * 60^{mins}), where SMV is the Standard Minute Value of the garment being produced. The SMV is the time industrial engineers estimate a fully efficient production line would take to produce one unit of the garment. When estimated to a common standard, the SMV thus allows us to compare the efficiency of production of different products – e.g., the efficiency of a line producing a tank top with an SMV of six minutes can be compared with the efficiency of a line producing a dress shirt with an SMV of 18 minutes.

We asked the factories to provide productivity records for each sewing line and day detailing on daily output, the number of defective units, the SMV of the garment being produced, the number of hours each line operated, and daily number of workers present and absent on the line. Not all factories record information on all of the variables. In some instances, the factories record data, but declined to provide it for certain outputs. For example, one factory declined to provide SMV data, and a few others do not have industrial engineering departments, and hence do not estimate SMVs by product. For other variables, there are sometimes differences in the specific data the factories record, though often these differences are not consequential. For example, for defects, we sometimes received defect rates (defective units / output) and sometimes the number of defective garments. Records on absenteeism would sometimes contain information on the numbers of workers assigned to the line, allowing to standardize the absenteeism numbers. At factories where this information was not included, we instead standardized the number of absent workers by the number of present workers provided in the productivity data.

In almost all factories, the three types of data (on productivity, defects, and absenteeism) was provided by different departments within the factories (usually the production, quality, and HR departments), and thus came in different formats, which required to enter the data separately and subsequently merge them to one document. Likewise, in most factories, the data we requested was provided in a digital format, usually a spreadsheet maintained by the factories, which allowed for easy collection and entering. At some factories, however, data was provided as copies of paper files, requiring the data be digitised. Ultimately, though, we harmonise the data so that variables are comparable across factories.

As we noted, the data from some factories did not contain the information necessary to calculate all of the outcomes of interest. This is especially the case for efficiency, where our standard calculation relies on the availability of the SMV data. Some of the factories that do not measure SMV have other data which can be used to estimate a roughly equivalent measure of efficiency. For example, four factories in the Phase 2 sample have information on daily targets for their sewing lines. By assuming that the targets are set such that line efficiency would be 100%, we can back out a 'synthetic SMV' by setting $\textit{Daily Target} * \textit{SMV} = \textit{workers} * \textit{daily hours} * 60^{mins}$.[1]

[1]Factories from which both SMV and targets are available show that targets are usually not set such that efficiency, in case the target is met, is 100%. Rather, efficiency in these cases would be around 50%, which is in line with the typical average efficiency in almost all Bangladeshi garment factories. Thus, the 'synthetic SMVs' which we back out using targets are likely to overstate actual SMVs by a factor of two. And indeed, efficiency values at those factories where we use 'synthetic SMVs' are on average twice as high as in the other factories (93% vs 47%). However, note that all analysis we conduct with the production data uses factory fixed effects, therefore relying only on

Table B.1: Production Data Availability

Outcome variable	Nbr. Factories
Productivity	14
Productivity, excluding synthetic SMV	10
Defects rate	16
Absenteeism	10
Productivity + Defects + Absenteeism	7
Productivity (excl. synth. SMV) + Defects + Absenteeism	4
Total Nbr. Factories with some Prod. Data	17

Notes: Table shows for how many factories participating in Phase 2, usable daily data on line-wise productivity, defects rates, and absenteeism could be collected.

From 17 of the 19 factories remaining in the project throughout, we were able to collect data for at least one of our three outcome variables of interest; productivity, defects, and absenteeism. Table B.1 shows from how many of these 17 factories we could collect enough data to construct each of the three variables, and for how many we could construct all three. While the availability of defects data is most widespread, productivity data is reduced by a number of factories recording neither SMVs nor targets. Finally, the availability of absenteeism data for our analysis is limited by a number of factories recording only daily absenteeism numbers for the whole factory (or sometimes the sewing floor), but not recording data on the sewing lines on which workers are assigned.

within factory variation in productivity. Given that for each factory we use either only productivity based on original or synthetic SMVs, the productivity data is consistent within each factory.

Appendix C

Worker Movement

While sewing worker are allocated to fixed lines, they do at times switch lines on a day-to-day basis to replace absent workers. With average daily absenteeism rates in the sample factories non-negligible at 3-5%, there is scope that within these reallocations, enough workers with relevant production knowledge on specific garments are moved across lines to drive the observed increases in productivity when other lines have already produced the same garment before. This would indicate that productivity spill-overs to later lines producing the same garment are driven by within-worker transfers of production knowledge across sewing lines, and not by knowledge exchange across workers.

To test for the likelihood that indeed short term movements of workers account for the higher productivity if other lines have already produced the garment before, a tentative test is conducted. If workers with experience on the garments are reallocated across lines, the lines from which the workers are taken should experience a negative effect on their productivity if they still produce the garment when further lines start producing the garment as well, and workers are reallocated away to these lines. Table C.1 below shows the results when regressing the daily efficiency of first lines that produced a garment in the factory on line chief - garment-day fixed effects, factory-week fixed effects, and on a dummy indicating that by that day, additional lines ('Additional Line') have also started producing the same garment. The results show that instead of a drop in productivity, if anything, the first line experiences an increase in productivity when other lines also start producing the garment. The source of these positive effects are not immediately clear. These effects could be due to reverse knowledge spill-over from the additional lines back to the first line, or due to other forms of peer effects, such as competition. However, these results are not in line with what could be expected if systematic movement of workers with production experience on certain garments would cause the observed productivity spill-over.

Table C.1: Worker Movement

VARIABLES	Efficiency
Additional Line	3.249***
	(0.838)
Constant	43.353***
	(3.640)
Observations	18,749
R^2	0.513
Factory-week FE	YES
L.Chief-Grmt.Day FE	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix D

Application of DiNardo et al. [1996]

The implementation of the reweighting approach of DiNardo et al. [1996] requires the estimation of two probit models; first, of a dummy indicating whether a unit i in the sample is selected for treatment ($T_i = 1$) on the unbalanced variable z_i, and, second, of a dummy indicating whether the unit is selected as control ($T_i = 0$) on z_i. The predicted probabilities $P(T = 1|z_i)$ for each unit i for being in the treatment group, and $P(T = 0|z_i)$ for being in the control group conditional on the unbalanced variable z, and the unconditional probabilities $P(T = 1)$ and $P(T = 0)$ of being selected into treatment or control sample, respectively, are then used to calculate weights w_i for each unit i according to:

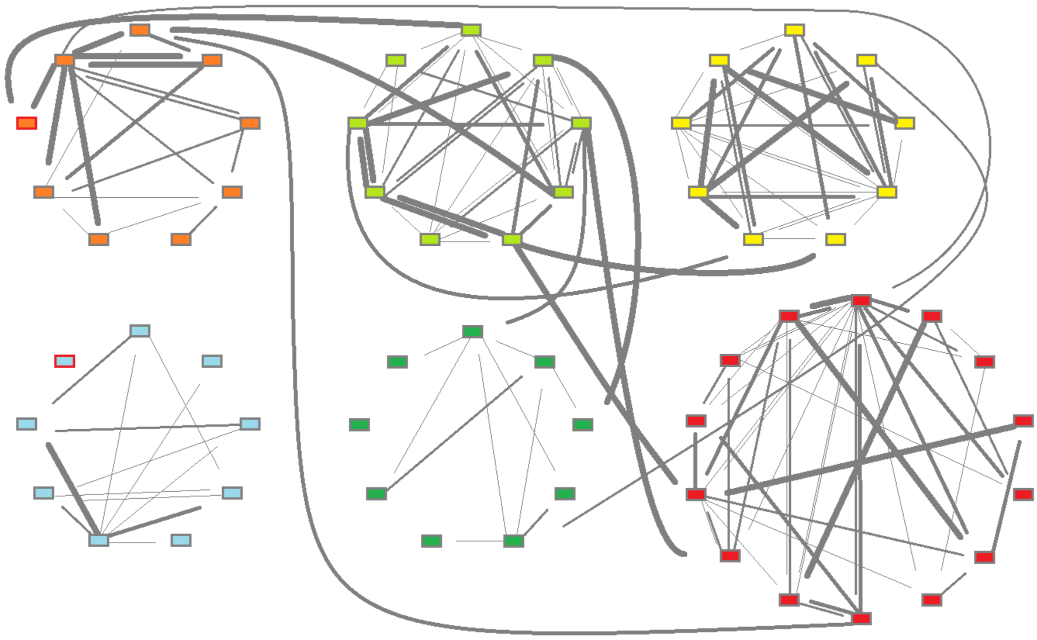
$$w_i = \frac{P(T = 0|z_i)P(T = 1)}{P(T = 1|z_i)P(T = 0)} \tag{D.1}$$

To implement the approach, I first regress, on a sample of all sewing lines, a dummy indicating that a sewing line is located on a floor selected for treatment, on the line's average efficiency on the first days it produced new garments that have already been produced on other lines before, during the pre-intervention time April and May 2014, controlling for factory fixed effects. The predicted values of this regression for each sewing line yield $P(T = 1|z_i)$ for calculating the weights w_i, according to equation D.1. Similarly, I also regress a dummy indicating that a line is located on a control floor on its average first-day efficiency during April and May 2014, to obtain $P(T = 0|z_i)$. I then reweight all observations from the treatment lines with weight w_i for the respective line (control units are not reweighted in this approach, therefore weights w_i for lines from control floors are set to 1).

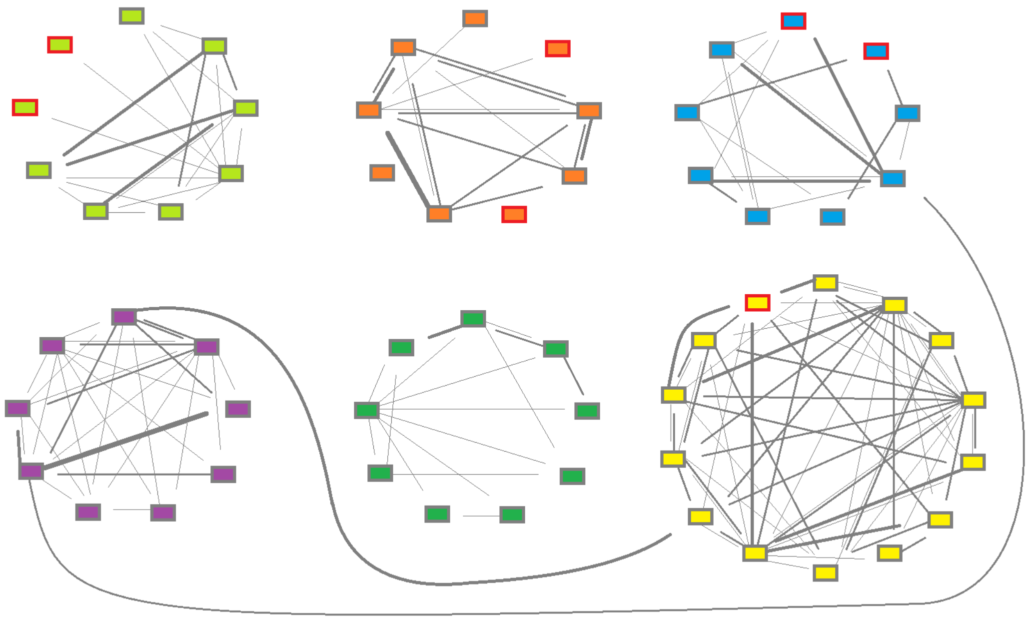
Appendix E

Network Graphs

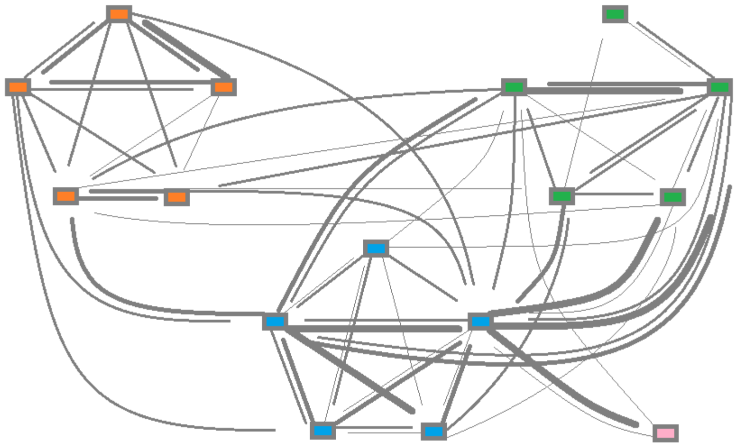
The six subsequent graphs depict the six network datasets collected at the factory-year level, two in 2013 and 2014 at Factory 1 and 2, respectively, and one collected at Factory 3 and 4 in 2014, respectively. Nodes with same filling colour represent line chiefs on the same floor. Thickness of links represents number of dimensions (visiting at home, spending lunch breaks together, being friends,...) that connections were reported on. Red frames of nodes represent line chiefs not surveyed due to being absent on the survey day. All graphs show directed links; the link is mentioned by the node which 'touches' the link, and to the node which does not touch the link.



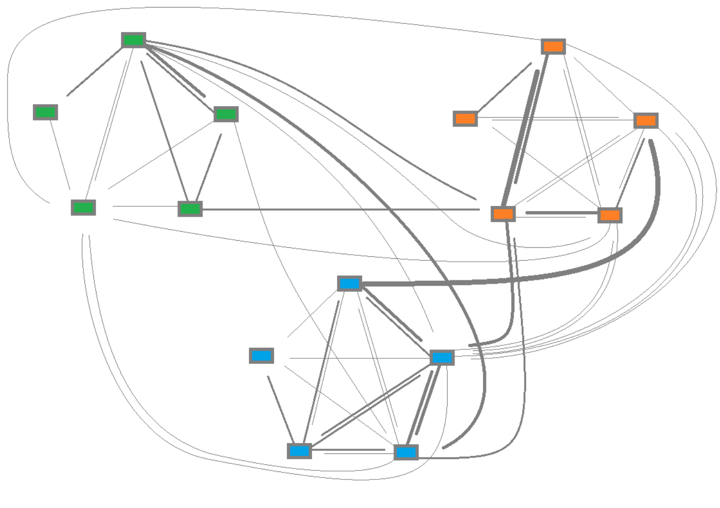
Factory 1, Year 2013



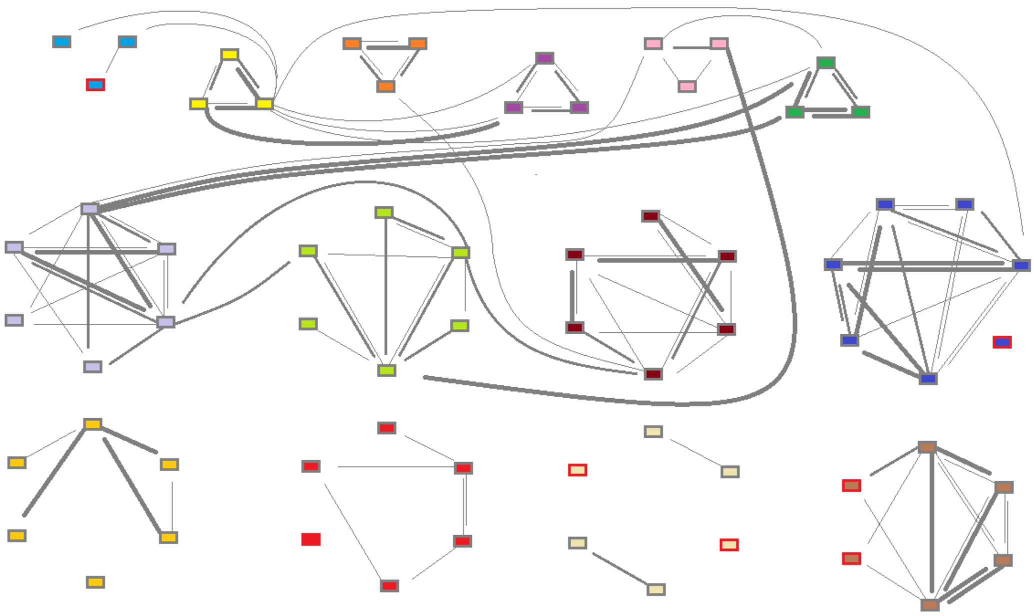
Factory 1, Year 2014



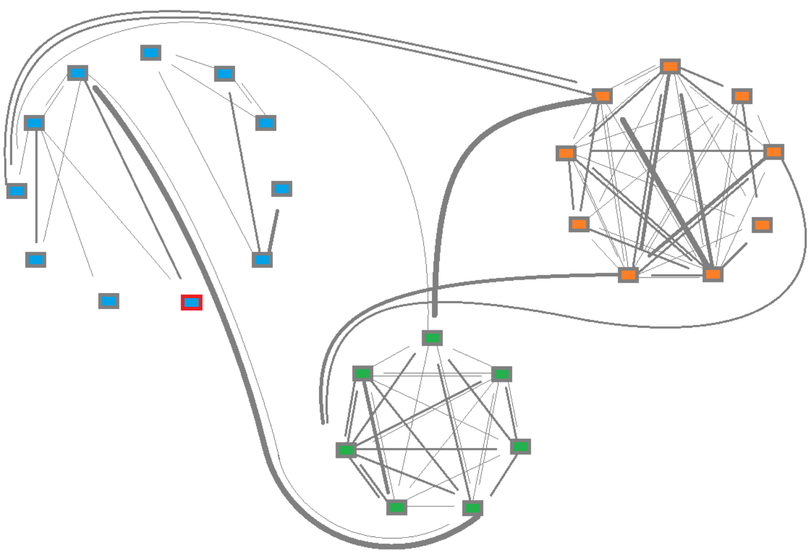
Factory 2, Year 2013



Factory 2, Year 2014



Factory 3 (Year 2014)



Factory 4 (Year 2014)

Bibliography

IGC Evidence Paper - Firms. 2014. URL `www.theigc.org/wp-content/uploads/2014/09/IGCEvidencePaperFirms.pdf`.

Adhvaryu Achyuta, Namrata Kala, and Anant Nyshadham. Management and Schocks to Worker Productivity: Evidence from Air Pollution Exposure in an Indian Garment Factory. Working Paper, University of Michigan, 2014.

Vivi Alatas, Abhijit Banerjee, Arun Chandrasekhar, Rema Hanna, and Benjamin Olken. Network Structure and the Aggregation of Information: Theory and Evidence from Indonesia. NBER Working Paper 18351, 2012.

Francesco Amodio and Miguel Martinez Carrasco. Input Allocation, Workforce Management and Productivity Spillovers: Evidence from Personnel Data. Working Paper, McGill University, 2015.

Kenneth Arrow. The Economic Implications of Learning by Doing. *The Review of Economic Studies*, 29(3):155–173, 1962.

David Atkin. Working for the Future: Female Factory Work and Child Health in Mexico. Working Paper, MIT, 2009.

David Atkin, Azam Chaudhry, Shamyla Chaudry, Amit Khandelwal, and Eric Verhoogen. Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan. NBER Working Paper 21417, 2015.

Matthew Backus. Why is Productivity Correlated with Competition? Working Paper, Columbia Business School, 2014.

Oriana Bandiera and Imran Rasul. Social Networks and Technology Adoption in Northern Mozambique. *Economic Journal*, 116(514):869–902, 2006.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Social Preferences and the Response to Incentives: Evidence from Personnel Data. *Quarterly Journal of Economics*, 120(3):917–962, 2005.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Social Incentives in the Workplace. *Review of Economic Studies*, 77(2):417–458, 2010.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Field Experiments with Firms. *Journal of Economic Perspectives*, 25(3):63–82, 2011.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Team Incentives: Evidence From A Firm Level Experiment. *Journal of the European Economic Association*, 11(5):1079–1114, 2013.

Oriana Bandiera, Andrea Prat, and Raffaella Sadun. Managing the Family Firm: Evidence from CEOs at Work. Technical report, C.E.P.R. Discussion Papers, 2015.

A. Banerjee, A. Chandrasekhar, E. Duflo, and M. Jackson. The Diffusion of Microfinance. *Science*, 341(6144), 2013.

Abhijit Banerjee and Esther Duflo. Growth Theory through the Lens of Development Economics. In Philippe Aghion and Steven Durlauf, editors, *Handbook of Economic Growth*. Elsevier, 1 edition, 2005.

Albert-Laszlo Barabasi and Reka Albert. Emergence of Scaling in Random Networks. *Science*, 286:509–512, 1999.

Lori Beaman. Social Networks and the Dynamics of Labor Market Outcomes: Evidence from Refugees Resettled in the U.S. *Review of Economic Studies*, 79(1): 128–161, 2012.

Lori Beaman, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. Powerful Women: Does Exposure Reduce Bias? *Quarterly Journal of Economics*, 124(4):1497–1540, 2009.

Lori Beaman, Ariel BenYishay, Jeremy Magruder, and Mushfiq Mobarak. Can Network Theory-based Targeting Increase Technology Adoption? Working Paper, Northwestern University, 2015.

Lanier Benkard. Learning and Forgetting: The Dynamics of Aircraft Production. *American Economic Review*, 90(4):1034–1054, 2000.

Andrew Bernard, Bradford Jensen, Stephen Redding, and Peter Schott. Firms in International Trade. *Journal of Economic Perspectives*, 21(3):105–130, 2007.

Marianne Bertrand and Kevin Hallock. The Gender Gap in Top Corporate Jobs. NBER Working Papers 7931, 2000.

Marianne Bertrand, Sendhil Mullainathan, and Erzo Luttmer. Network Effects and Welfare Culture. *Quarterly Journal of Economics*, 115(3):1019–1055, 2000.

Marianne Bertrand, Sandra Black, Sissel Jensen, and Adriana Lleras-Muney. Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labor Market Outcomes in Norway. NHH Dept. of Economics Discussion Paper No. 28/2014, 2014.

Tulikaa Bhatia, Puneet Manchanda, and Harikesh Nair. Asymmetric Per Effects in Physician Prescription Behavior: The Role of Opinion Leaders. Stanford University School of Business Working Paper 1970, 2006.

Nicholas Bloom and John Van Reenen. Measuring and Explaining Management Practices Across Firms and Countries. *Quarterly Journal of Economics*, 122(4): 1351–1408, 2007.

Nicholas Bloom, Raffaella Sadun, and John Van Reenen. The Organization of Firms across Countries. *Quarterly Journal of Economics*, 127(4):1663–1705, 2012.

Nicholas Bloom, Ben Eifert, David McKenzie, Aprajit Mahajan, and John Roberts. Does Management Matter: Evidence from India. *Quarterly Journal of Economics*, 128(1):1–51, 2013.

Nicholas Bloom, Kalina Manova, John Van Reenen, and Zhihong Yu. Management, Product Quality and Trade: Evidence from China. Work in Progress, Stanford University, 2015.

Miriam Bruhn, Dean Karlan, and Antoinette Schoar. The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico. C.E.P.R. Discussion Paper 8887, 2012.

Ronald Burt. Decay Functions. *Social Networks*, 22(1):1–28, 2000.

Jing Cai, Alain De Janvry, and Elisabeth Sadoulet. Social Networks and the Decision to Insure. *American Economic Journal: Applied Economics*, 7(2):81–108, 2015.

Antoni Calvo-Armengol, Eleonora Patacchini, and Yves Zenou. Peer Effects and Social Networks in Education. *Review of Economic Studies*, 76(4):1239–1267, 2009.

Arun Chandrasekhar. Econometrics of Network Formation. In Yann Bramoulle, Andrea Galeotti, and Brian Rogers, editors, *Oxford Handbook on the Economics of Networks*. 2015.

Arun Chandrasekhar and Matthew Jackson. Tractable and Consistent Random Graph Models. NBER Working Paper 20276, 2014.

Arun Chandrasekhar, Cyntia Kinnan, and Horacio Larreguy. Social Networks as Contract Enforcement: Evidence from a Lab Experiment in the Field. NBER Working Paper 20259, 2014.

Raghabendra Chattopadhyay and Esther Duflo. Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica*, 72(5):1409–1443, 2004.

Andrew Clark and Youenn Loheac. 'It Wasn't Me, It Was Them', Social Influence and Risky Behaviour by Adolescents. *Journal of Health Economics*, 26(4):763–784, 2007.

James Coleman, Elihu Katz, and Herbert Menzel. *Medical Innovation: A Diffusion Study*. Bob-Mills, Indianapolis, Ind., 1966.

Margeritha Comola. Estimating Local Network Externalities. Working Paper, Paris School of Economics, 2012.

Timothy Conley and Christopher Udry. Learning about a new Technology: Pineapple in Ghana. *American Economic Review*, 100(1):35–69, 2010.

Jonathon Cummings and Robert Cross. Structural Properties of Work Groups and their Consequences for Performance. *Social Networks*, 25(3):197–210, 2003.

Sergio Currarini, Matthew Jackson, and Paolo Pin. An Econometric Model of Friendship: Homophily, Minorities and Segregation. *Econometrica*, 77(4):1003–1045, 2009.

Sanghamitra Das, Kala Krishna, Rohini Somanathan, and Sergey Lychagin. Back on the Rails: Competition and Productivity in State-owned Industry. *American Economic Journal: Applied Economics*, 5(1):136–62, 2013.

Joachim De Weerdt and Stefan Dercon. Risk-Sharing Networks and Insurance against Illness. *Journal of Development Economics*, 81(2):337–356, 2006.

Joachim De Weerdt and Marcel Fafchamps. Social Identity and the Formation of Health Insurance Networks. *Journal of Development Studies*, 47(8):1152–1177, 2011.

Bella DePaulo and Jeffrey Fisher. The Costs of Asking for Help. *Basic and Applied Social Psychology*, 1(1):23–35, 1980.

Cristian Dezsö and David Gaddis Ross. Does Female Representation in Top Management improve Firm Performance? A Panel Data Investigation. *Strategic Management Journal*, 33(9):1072 – 1089, 2012.

John DiNardo, Nicole Fortin, and Thomas Lemieux. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semi-Parametric Approach. *Econometrica*, 64(5):1001–1044, 1996.

Esther Duflo and Emmanuel Saez. The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment. *Quarterly Journal of Economics*, 118(3):815–842, 2003.

Esther Duflo, Michael Greenstone, Rohini Pande, and Nicholas Ryan. Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India. *Quarterly Journal of Economics*, 128(4):1499 – 1545, 2013.

Paul Erdos and Alfred Renyi. On Random Graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

Marcel Fafchamps and Flore Gubert. Risk Sharing and Network Formation. *American Economic Review*, 97(2):75–79, 2007a.

Marcel Fafchamps and Flore Gubert. The Formation of Risk Sharing Networks. *Journal of Development Economics*, 83(2):326–350, 2007b.

Katherine Faust. Very local Structure in Social Networks. *Sociological Methodology*, 37(1):209–256, 2007.

Thomas Feeley and George Barnett. Predicting Employee Turnover from Communication Networks. *Human Communication Research*, 23:370–387, 1997.

Thomas Feeley, Jennie Hwang, and George Barnett. Predicting Employee Turnover from Friendship Networks. *Journal of Applied Communication Research*, 36:56–73, 2008.

Andrew Foster and Mark Rosenzweig. Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy*, 103(6):1176–1209, 1995.

Lucia Foster, John Haltiwanger, and Chad Syverson. Reallocation, Firm Turnover and Efficiency: Selection on Productivity or Profitability. *American Economic Review*, 1(98), 2008.

Stephan Fuchs. The Stratified Order of Gossip: Informal Communication in Organizations and Science. *Soziale Systeme*, 1(1):44, 1995.

Clifford Geertz. The Bazaar Economy: Information and Search in Peasant Marketing. *American Economic Review*, 68:28–32, 1978.

Robert Gibbons and Rebecca Henderson. Relational Contracts and Organizational Capabilities. *Organization Science*, 23(5):1350–1364, 2012.

Dylan Glover, Amanda Pallais, and William Pariente. Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores. Working Paper, Harvard, 2015.

Uri Gneezy and Aldo Rustichini. Gender and Competition at a Young Age. *American Economic Review*, 94(2):377–381, 2004.

Paul Goldsmith-Pinkham and Guido Imbens. Social Networks and the Identification of Peer Effects. *Journal of Business, Economics and Statistics*, 31(3):253–264, 2013.

Bryan Graham. Methods of Identification in Social Networks. *Annual Review of Economics*, 7:465–485, 2015a.

Bryan Graham. An Econometric Model of Link Formation with Degree Heterogeneity. Working Paper, University of California, Berkely, 2015b.

Mark Granovetter. The Strength of Weak Ties. *American Journal of Sociology*, 78: 1360–1380, 1973.

Mark Granovetter. *Getting a Job: A Study of Contacts and Careers*. University of Chicago Press, Chicago, 2.Ed, 1995.

Barton Hamilton, Jack Nickerson, and Hideo Owan. Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams in Productivity and Participation. *Journal of Political Economy*, 111(3):465–497, 2003.

Rachel Heath. Why do Firms Hire using Referrals? Evidence from Bangladeshi Garment Factories. Working Paper, University of Washington, 2015.

Rachel Heath and Mushfiq Mobarak. Manufacturing growth and the lives of Bangladeshi women. *Journal of Development Economics*, 115:1–15, 2015.

Igal Hendel and Yossi Spiegel. Small Steps for Worker, a Giant Leap for Productivity. *American Economic Journal: Applied Economics*, 6(1):73–90, 2014.

Jonas Hjort. Ethnic Divisions and Production in Firms. *Quarterly Journal of Economics*, 129(4):1899–1946, 2014.

Harrison Hong, Jeffrey Kubik, and Jeremy Stein. Social Interaction and Stock Market Participation. *Journal of Finance*, 59(1):137–163, 2004.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448, 2009.

Casey Ichniowski, Kathryn Shaw, and Giovanna Prennushi. The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines. *American Economic Review*, 87(3):291–313, 1997.

Yannis Ioannides and Linda Datcher Loury. Job Information Networks, Neighborhood Effects, and Inequality. *Journal of Economic Literature*, 42(4):1056–1093, 2004.

Matthew Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, NJ, 2008.

Matthew Jackson and Brian Rogers. Meeting Strangers and Friends of Friends: How Random are Social Networks. *American Economic Review*, 97(3):890–915, 2007.

Matthew Jackson and Asher Wolinsky. A Strategic Model of Social and Economic Networks. *Journal of Economic Theory*, 71(1):44–74, 1996.

Denise Kandel. Homophily, Selection, and Socialization in Adolescent Friendships. *American Journal of Sociology*, 84(2):427–436, 1978.

Takao Kato and Pian Shu. Competition, Group Identity, and Social Networks in the Workplace: Evidence from a Chinese Textile Firm. IZA Discussion Paper 6219, 2011.

Christopher Ksoll, Rocco Macchiavello, and Ameet Morjaria. The Effect of Ethnic Violence on an Export-Oriented Industry. C.E.P.R. Discussion Papers 8074, 2010.

Paul Lazarsfeld and Robert Merton. Friendship as a Social Process: A Substantive and Methodological Analysis. In M. Berger, editor, *Freedom and Control in Modern Society*. Van Nostrand, New York, 1954.

Paul Lazarsfeld, Bernard Berelson, and Hazel Gaudet. *The People's Choice: How the Voters Make up his Mind in a Presidential Campaign*. Columbia University Press, New York, 1944.

Edward Lazear. Performance Pay and Productivity. *American Economic Review*, 90(5):1346–1361, 2000.

Fiona Lee. The Social Costs of Seeking Help. *The Journal of Applied Behavioral Science*, 38(1):17–35, 2002.

Stephen Leider, Markus Mobius, Tanya Rosenblat, and Quoc-Anh Do. Directed Altruism and Enforced Reciprocity in Social Networks. *Quarterly Journal of Economics*, 124(4):1815–1851, 2009.

Steve Levitt, John List, and Chad Syverson. Towards and Understanding of Learning by Doing: Evidence from an Automobile Plant. *Journal of Political Economy*, 121(4):643–681, 2013.

Robert Lucas. Making a Miracle. *Econometrica*, 61(2):251–272, 1993.

Alexandre Mas and Enrico Moretti. Peers at Work. *American Economic Review*, 99(1):112–145, 2009.

David Matsa and Amalia Miller. A Female Style in Corporate Leadership? Evidence from Quotas. *American Economic Journal: Applied Economics*, 5(3):136–69, 2013.

David McKenzie and Christopher Woodruff. Business Practices in Small Firms in Developing Countries. NBER Working Paper 21505, 2015.

McKinsey. Bangladesh's ready made garments landscape: The challange of growth. Technical report, McKinsey&Company, Apparel, Fashion & Luxury Practice, 2011.

John McMillan and Christopher Woodruff. Interfirm Relationships and Informal Credit in Vietnam. *Quarterly Journal of Economics*, 114(4):1285–1320, 1999.

Angelo Mele. A Structural Model of Segregation in Social Networks. Working Paper, Johns Hopkins University, 2013.

Stanley Milgram. The Small-World Problem. *Psychology Today*, 2:60–67, 1967.

Kevin Mossholder, Randall Settoon, and Stephanie Henagan. A Relational Perspective on Turnover: Examining Structural, Attitudinal, and Behavioral Predictors. *Academy of Management Journal*, 48:607–618, 2005.

Kaivan Munshi. Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *Quarterly Journal of Economics*, 118(2):549–597, 2003.

Kaivan Munshi. Social Learning in a heterogeneous population: Technology Diffusion in the Indian Green Revolution. *Journal of Development Economics*, 73: 185–213, 2004.

Kaivan Munshi. Strength in Numbers: Networks as a Solution to Occupational Traps. *Review of Economic Studies*, 78:1069–1101, 2011.

Kaivan Munshi and Mark Rosenzweig. Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. Working Paper, Cambridge University, 2013.

Charles Myers and George Schultz. *The Dynamics of a Labor Market.* Prentice-Hall, New York, 1951.

Daniel Nagin, James Rebitzer, Seth Sanders, and Lowell Taylor. Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment. *American Economic Review*, 92(4):850–873, 2002.

Muriel Niederle and Lise Vesterlund. Do Women Shy Away From Competition? Do Men Compete Too Much? *Quarterly Journal of Economics*, 122(3):1067–1101, 2007.

Muriel Niederle, Carmit Segal, and Lise Vesterlund. How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness. *Management Science*, 59(1):1–16, 2013.

Dotan Persitz. Power in the Heterogeneous Connections Model: The Emergence of Core-Periphery Networks. Fondazione Eni Enrico Mattei Working Paper 2009.042, 2009.

Steven Ruggles, Trent Alexander, Katie Genadek, Ronald Goeken, Matthew Schroeder, and Matthew Sobek. Integrated public use microdata series: Version 5.0 [machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2010.

Antoinette Schoar. The Importance of Being Nice: Evidence from a Supervisory Training Program in Cambodia. mimeo, MIT, 2011.

Lan Shi. Incentive Effects of Piece Rate Contracts: Evidence from two small Field Experiments. *BE Journal of Economic Analysis and Policy*, 10(1):Article 61, 2010.

Alan Sorensen. Social Learning and Health Plan Choice. 2006.

Chad Syverson. Product Substitutability and Productivity Dispersion. *Review of Economics and Statistics*.

Peter Thompson. Learning by Doing. In B. Hall and N. Rosenberg, editors, *Handbook of Economics of Technical Change*. Elsevier, North-Holland, 2010.

Peter Thompson and Rachel Thornton. Learning from Experience and Learning from Others: An Exploration of Learning and Spill-overs in Wartime Shipbuilding. *American Economic Review*, 91(5):1350–1368, 2001.

Brian Uzzi. The Sources and Consequences of Embeddedness for the Economic Performance of an Organization: The Network Effect. *American Sociological Review*, 61:674–698, 1996.

Gerard Weisbuch, Alan Kirman, and Dorothea Herreiner. Market Organization. Discussion Paper, University of Bonn, 1996.

Gerard Weisbuch, Alan Kirman, and Dorothea Herreiner. Market organization and trading relationships. *Economic Journal*, 110(463):411–436, 2000.

Theodore Wright. Factors affecting the Cost of Airplanes. *Journal of Aeronautical Sciences*, 3(4):122–128, 1936.